Therapeutic Area Data Standards User Guide for Alzheimer's Disease and Mild Cognitive Impairment

Version 2.0

Prepared by the

CFAST Alzheimer's Development Team

Notes to Readers

- This is version 2.0 of the Therapeutic Area Data Standards User Guide for Alzheimer's Disease and Mild Cognitive Impairment. It includes information about research concepts used in Alzheimer's and Mild Cognitive Impairment trials. It also includes information about those concepts.
- This document corresponds to the SDTM v1.4 and SDTMIG v3.2.
- The TAUG-Alzheimer's v2.0 package includes a user guide (this document) and five draft domains.

Revision History

Date	Version	Summary of Changes
2013-12-16	2.0	TAUG-Alzheimer's v2.0
2013-10-04	2.0 Draft	Draft for public review
2013-08-26	2.0 Draft	Draft for internal review with updated examples from additional domains
2011-09-09	1.0	Alzheimer's Disease User Guide v1.0 release

See Appendix D for Representations and Warranties, Limitations of Liability, and Disclaimers.

CONTENTS

1 Introduction

This CDISC Therapeutic Area Data Standards User Guide for Alzheimer's Disease and Mild Cognitive Impairment (TAUG-Alzheimer's) was developed under the Coalition for Accelerating Standards and Therapies (CFAST) initiative.

CFAST, a joint initiative of CDISC and the Critical Path Institute (C-Path), was launched to accelerate clinical research and medical product development by facilitating the establishment and maintenance of data standards, tools, and methods for conducting research in therapeutic areas important to public health. CFAST partners include TransCelerate BioPharma Inc. (TCB), the United States Food and Drug Administration (FDA), and the National Cancer Institute Enterprise Vocabulary Services (NCI EVS), with participation and input from many other organizations.

1.1 Purpose

This document comprises Version 2.0 (v2.0) of the Alzheimer's disease-specific Therapeutic Area User Guide (TAUG-Alzheimer's) to be used as a supplement to the CDISC Study Data Tabulation Model (SDTM) Implementation Guide for Human Clinical Trials (SDTMIG). This user guide was prepared by the CFAST Alzheimer's Development Team with participation from the Critical Path Institute's Coalition Against Major Diseases (CAMD), and with input from the Alzheimer's Disease Neuroimaging Initiative (ADNI). It is intended to guide the organization, structure, and format of standard Alzheimer's disease (AD) and mild cognitive impairment (MCI) clinical trial tabulation datasets that can be used to support data review and analysis and submitted to a regulatory authority such as the FDA.

With regard to clinical trials of AD and MCI, this guide describes a specific implementation of a subset of the domains whose general implementation is described in the current version of the SDTMIG. This document also makes use of newly drafted SDTM-based domains, and shows rules for and examples of implementing these domains specifically for trials of Alzheimer's disease or mild cognitive impairment.

This document does not replace, supersede, nor otherwise override any rules or requirements of the current SDTM and SDTMIG. Knowledge of this document alone is not a substitute for knowledge of SDTM nor is it sufficient to produce complete, SDTM-compliant regulatory submissions of Alzheimer's or MCI clinical trials data. The TAUG-Alzheimer's v2.0 should be used in close concert with the most current version of the SDTMIG and the most current version of the CDISC Study Data Tabulation Model (available at http://www.cdisc.org/standards). The SDTM describes the general model for representing clinical study data that is submitted to regulatory authorities and should be read prior to reading the SDTMIG. An understanding of both of these documents is required before attempting to read and understand the TAUG-Alzheimer's v2.0.

This document is intended for companies and individuals involved in the collection, preparation, and analysis of clinical data that will be submitted to regulatory authorities.

Domains for which there are no current Alzheimer's-specific rules are not included in this document. For information on these domains, refer to the current version of the SDTM/SDTMIG and other CDISC supplemental guides.

1.2 Organization of this Document

This TAUG-Alzheimer's v2.0 uses a transitional format, between 1) that of the legacy therapeutic area user guides that mirrored the SDTMIG in their organization, and 2) the style of therapeutic area user guides developed under the CFAST initiative and piloted by the TAUG-Asthma, which have a more clinically oriented organization.

- Section 1, Introduction, provides an overall introduction to the purpose and goals of the Alzheimer's disease project.
- Section 2, Subject and Disease Characteristics, covers data that are usually collected once at the beginning of a study.
- Section 3, Disease Assessments, covers data that are used to evaluate disease severity, control, or progression. These are usually collected repeatedly during a study, and are often used as efficacy endpoints.
- Appendices provide additional background material and describe other supplemental material relevant to Alzheimer's disease.

A list of domains used in the examples in this document, and the sections in which they appear, are given below. The examples in this user guide, which include devices domains, record ancillary devices only. Unpublished draft domains are maintained on the CDISC Portal under *Teams Projects/ SDTM/ Working Documents for Production Release*.

Domains from SDTMIG	Section(s)
Interventions	
AG – Procedure Agents*	3.2.1.2, 3.2.2.2
PR – Procedures	3.1.2, 3.2.1.2, 3.2.2.2
Events	
MH – Medical History	2.2.1, 2.3.1 (APMH)
Findings	
LB – Laboratory Test Results	3.1.2
MO – Morphology	3.2.1.2
NV –Nervous Systems Findings*	3.2.2.2

** Domain was not published in SDTMIG 3.2 and is not final.*

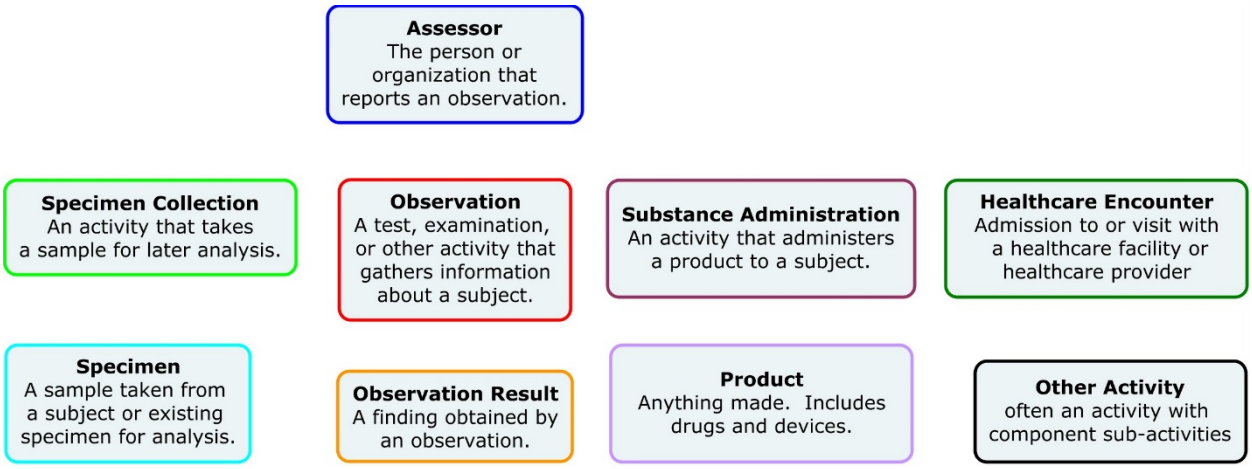
Domains from SDTMIG-PGx*	Section(s)
Biospecimen Domains	
BE – Biospecimen Event	3.1.2
BS – Biospecimen	3.1.2
Pharmacogenomics/Genetics Domains	
PF – Pharmacogenomics Findings	2.1.1

**Implementation guide is not final.*

Domains from SDTMIG-MD	Section(s)
DI – Device Identifier	3.1.2, 3.2.1.2, 3.2.2.2
DU – Device In-Use	3.1.2, 3.2.1.2, 3.2.2.2
DO – Device Properties	3.1.2, 3.2.1.2, 3.2.2.2

1.3 Concept Maps

This document uses concept maps to explain clinical processes and research concepts. Concept maps, also sometimes called mind maps, are diagrams which include "bubbles" representing concepts/ideas/things and labeled arrows that represent the relationships between the concepts/ideas/things. They are generally easier to draw and more accessible than more formal modeling diagrams, such as Unified Modeling Language (UML) diagrams.



Color Coding Key for Concept Maps

The diagrams in this document use the color-coding shown above. The classification of concepts represented by these colors is based on classes in the BRIDG model. These colors have been used to highlight concepts that occur commonly in clinical data, and therefore give rise to common patterns of data.

1.4 Controlled Terminology

Terminology applicable to CDASH data collection fields is either in production or under development by the CDISC Terminology Team at the time of publication of this document. Production terminology is published by the National Cancer Institute's Enterprise Vocabulary Services (NCI EVS) and is available at: http://www.cancer.gov/cancertopics/cancerlibrary/terminologyresources/cdisc.

CDISC Controlled Terminology is updated quarterly. Because this document is a static publication, it refers readers to the NCI EVS page for CDISC terminology (at the link given above). For the same reason, this document cannot claim to use controlled terminology in either the lists of laboratory tests or in the examples provided; users should not refer to these as the ultimate authority on what terms to use.

1.5 Relationships to Other Standards

This section describes the relationship of this document to other standards, whether CDISC or external. This document does not replace the foundational CDISC standards or their implementation guides. The user should read those standards and implementation guides before applying the advice in this user guide.

This TAUG-Alzheimer's v2.0 is intended to replace the Alzheimer's disease SDTM User Guide v1.0. The examples contained in this document were developed to expand upon work published in the SDTMUG-Alzheimer's v1.0 in order to make the guide more useful in clinical studies, with a particular emphasis on early-disease assessments.

The primary sources of inputs for new implementations of the standard shown in this version were the Alzheimer's disease Neuroimaging Initiative (ADNI, http://www.adni-info.org/) and the biomarkers working group of the Coalition Against Major Diseases (CAMD), a program of the Critical Path Institute (www.c-path.org). Of the data specified by these two sources, the scope was limited to three major areas: clinical scales of cognition/function,

imaging biomarkers, and cerebrospinal fluid (CSF) biomarkers. Note that all clinical scales implemented for Alzheimer's disease are maintained as standalone supplements maintained on the CDISC website. See Section 3.3 for details.

This document uses domains and assumptions that are not final at the time of publication, and are therefore subject to change or deletion without formal notice. Please check the most recent versions of the SDTM and SDTMIG to ascertain their current status.

- The Nervous System Findings (NV) domain is still a work in progress (i.e., a "draft" domain) that has not appeared prior to this document.
- The Procedure Agents (AG) domain is a draft domain which first appeared as a part of the TAUG-Asthma v1.0 package.
- The Biospecimen (BS) and Biospecimen Event (BE) domains are draft domains expected to be released as part of the SDTM Implementation Guide for Pharmacogenomics/Genetics v1.0.
- The Procedures (PR) domain has been through public review and will be released with SDTM v1.4 and SDTMIG v3.2.
- The Morphology (MO) domain has been through public review and will be released with SDTM v1.4 and SDTMIG v3.2.
- The Pharmacogenomics Findings (PF) domain has appeared previously as a part of the Virology Therapeutic Area Data Standards User Guide v1.0, but is still a work in progress. The most recent version of this domain to date is included in this package.

In some cases when a definitive SDTM modeling approach does not exist, a suggested approach is offered but may be subject to change over time.

There are multiple types of data that have existing CDASH and SDTM standards that can be used in Alzheimer's disease studies without additional development or customization, such as Demographics, Adverse Events, Clinical Events, Concomitant Medications, etc. Domains whose uses are not explicitly represented in the examples within this document can still be used according to the rules of those domains as needed. Refer to the current SDTMIG for implementation examples using those domains.

1.5.1 Summary of Changes from Previous Version

- Added examples detailing the data modeling of assays and protocol parameters used in the generation of the CSF biomarkers endpoints shown in Version 1.0 Laboratory Results (LB). Those examples are still shown in Version 2.0, but now have detailed sample collection and handling parameters represented in various other domains as well.
- Added examples of imaging biomarkers including the actual endpoints (volumetric, Standard Uptake Value Ratios) and the imaging protocol parameters used to generate these endpoints (via MRI, PET, and PET/CT).
- Added ten additional clinical scales relevant to AD and MCI to the Questionnaire Supplements to the SDTMIG maintained as separate documents on the CDISC website.
- Demographics domain (DM) examples were removed. There are no AD-specific demographics rules that should be standardized across studies. Sponsors should refer to the DM domain in the SDTMIG for details on creating this dataset.
- Family history was moved from Medical History (MH) to the Associated Persons Medical History (APMH) Domain.
- ApoE genetics data examples were moved from the Subject Characteristics (SC) domain to the Pharmacogenomics Findings domain (PF).
- The terminology for handling "Date of Onset of Cognitive Problem" in the Supplemental Qualifiers domain (SUPPMH) was changed to match other published guides.

1.6 Known Issues

- Additional assumptions for Alzheimer's-specific implementation of the domain examples demonstrated in this guide are contained in Appendix C and these assumptions should be applied when representing domain data for Alzheimer's studies. Since Therapeutic Area User Guides are not arranged by domain like the SDTMIG, examples of domains may appear in multiple sections together with the concepts covered in those sections, thus precluding a single, logical placement for assumptions for general domains that reappear in the TAUG. CDISC recognizes that including these in an appendix this is not ideal. This may be addressed in future releases.
- Software analysis is represented as an analytical method in the imaging examples contained in the PET and MRI sections, and the name and version of the software is represented in Supplemental Qualifiers. Some regard this software as a standalone device. Therefore, it may be represented in the device domains in future releases pending CDISC SDS review and decision.
- CDISC SRC has recommended that only CDASH-compliant CRF examples be shown in CDISC guides. The example CRF in the Associated Persons Medical History example is not CDASH-compliant. This will be addressed in a future release."
- Users should refer to SDTMIG section 4.1.2.4 for rules and recommendations on use of case for text submitted in data. Some values shown in the example values do not strictly adhere to these recommendations, particularly for --TRT, --TERM, --ORRES, DIVAL, and QVAL variables.

2 Subject and Disease Characteristics

2.1 Genetics of Alzheimer's Disease

Specific variants of the apolipoprotein E (ApoE) gene have been implicated in the development and progression of Alzheimer's. As with any gene, an individual with a normal number of chromosomes (46) would have two copies (alleles) of the ApoE gene. There are three possible alleles of the ApoE gene, numbered as "2," "3" and "4." As such, investigators are primarily interested in a categorical genotype notation that indicates which of three possible isoforms a subject carries on each of his or her two alleles. Determination of genotype may be done by a variety of genetic testing methods, such as restriction fragment length polymorphism (RFLP) analysis. The results come in the form of a numerical representation of each of the two alleles a subject carries, giving the following possible six categorical combinations: "2-2", "2-3", "2-4", "3-3", "3-4" or "4-4". Subjects who carry two copies of isoform 4 ("4-4") are widely regarded to have a predisposition to developing Alzheimer's and showing faster disease progression. Isoform 3 is considered neutral whereas isoform 2 may be protective against Alzheimer's, but this has not yet been confirmed. Ongoing research seeks to understand the role and impact that the various combinations of these isoforms have on disease course. Subjects who are heterozygous for any given isoform (e.g., subjects who carry only one copy of isoform 4) may have a different disease course than subjects who are homozygous for any isoform (e.g., subjects who carry isoform 4 on both alleles), therefore it is important to record the results of each allele as a separate record as opposed to concatenating both allele results in one record, to allow for easier sorting of each allele independently

This section explains how to represent the various combinations of isoforms that may be coded on each allele as determined by genetic testing.

2.1.1 Examples for the Genetics of Alzheimer's Disease

Example
This example shows the genotyping result for the Apolipoprotein E (ApoE) gene as two allele results for two subjects. Genotyping information is represented in the Pharmacogenomics Findings (PF) domain.

Rows 1-2: Show how to represent the results of a restriction fragment length polymorphism (RFLP) analysis (shown in PFMETHOD) used to make the determination of the isoforms represented by each allele of the Apolipoprotein E gene present in the subject. PFGENRI shows that the region of interest for this analysis was the Apolipoprotein E gene.
Rows 3-4: Show the results of the same analysis for another subject.

pf.xpt

Row	STUDYID	DOMAIN	USUBJID	PFSEQ	PFTESTCD	PFTEST	PFGENTYP	PFGENRI	PFCAT	PFORRES	PFSTRESC	PFMETHOD	PFBLFL	VISITNUM	PFDTC
1	ABC123	PF	AD01-101	1	ALE	Allele	GENE	ApoE	GENE EXPRESSION	3	3	RFLP ANALYSIS	Y	1	2013-05-22
2	ABC123	PF	AD01-101	2	ALE	Allele	GENE	ApoE	GENE EXPRESSION	4	4	RFLP ANALYSIS	Y	1	2013-05-22
3	ABC123	PF	AD01-102	1	ALE	Allele	GENE	ApoE	GENE EXPRESSION	4	4	RFLP ANALYSIS	Y	1	2013-05-22
4	ABC123	PF	AD01-102	2	ALE	Allele	GENE	ApoE	GENE EXPRESSION	4	4	RFLP ANALYSIS	Y	1	2013-05-22

2.2 Medical History

The following section describes how to represent medical history, both for Alzheimer's/MCI as well as general medical history terms. Note that neither MHTERM nor MHCAT are subject to controlled terminology. These examples are intended to highlight how the primary diagnosis may be categorized differently from general medical history, and how using generalized MHTERM values for Alzheimer's disease and Mild Cognitive Impairment is useful when multiple MHDECOD values may result from coding these terms to standard dictionaries. Sponsors should decide if they intend to use such dictionaries, and then decide which terms apply to their subjects' diagnostic criteria.

Since formal diagnosis of Alzheimer's is challenging, many investigators are also interested in collecting date of onset of symptoms. Reporting this onset date is also shown below.

2.2.1 Examples for Medical History

For additional assumptions, see Appendix C.

Example 1
Below are specific Alzheimer's/MCI Medical History examples for representing:

- General medical history entered as free text
- Primary diagnosis information

Rows 1, 4, 7: MHCAT shows the PRIMARY DIAGNOSIS for three separate subjects, one with Alzheimer's disease and two with mild cognitive impairment. Note that in this case MHDECOD is not populated.
Rows 2-3, 5-6: Indicate general medical history with MHCAT=GENERAL

mh.xpt

Row	STUDYID	DOMAIN	USUBJID	MHSEQ	MHTERM	MHDECOD	MHCAT	MHSTDTC
1	ABC123	MH	AD01-101	1	Mild Cognitive Impairment		PRIMARY DIAGNOSIS	2001-05
2	ABC123	MH	AD01-101	2	Blind Left Eye	Blindness unilateral	GENERAL	1999-07-07
3	ABC123	MH	AD01-101	3	Traumatic Brain Injury	Traumatic brain injury	GENERAL	2005-03-28
4	ABC123	MH	AD01-102	1	Alzheimer's Disease		PRIMARY DIAGNOSIS	2003-05
5	ABC123	MH	AD01-102	2	Partial Gastrectomy	Gastrectomy	GENERAL	2006-06-21
6	ABC123	MH	AD01-102	3	Arterial Hypertension	Hypertension	GENERAL	2006-07-26
7	ABC123	MH	AD01-103	1	Mild Cognitive Impairment		PRIMARY DIAGNOSIS	2002-07

Example 2: Onset of Symptoms
There is some controversy over the method of recording the onset of symptoms. This example uses Supplemental Qualifiers, but this approach is under evaluation by the SDS team and may be revised in the future.

The onset of symptoms date (QNAM=MHOSDTC) is when the symptoms of Alzheimer's disease or mild cognitive impairment were first observed by the subject or a family member. The actual diagnosis date (MHSTDTC) is the date the subject's physician officially diagnosed the disease. Where collected, onset date of symptoms (for either Alzheimer's or MCI) should be mapped to SUPPMH as in the following example. According to precedent, QNAM is MHOSDTC and QLABEL is "Onset of Symptoms Date".

suppmh.xpt

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL	QORIG	QEVAL
1	ABC123	MH	AD01-101	MHSEQ	1	MHOSDTC	Onset of Symptoms Date	1999-05	CRF	MOTHER
2	ABC123	MH	AD01-102	MHSEQ	1	MHOSDTC	Onset of Symptoms Date	2000-02	CRF	FATHER
3	ABC123	MH	AD01-103	MHSEQ	1	MHOSDTC	Onset of Symptoms Date	1998	CRF	STUDY SUBJECT

2.3 Family History

Understanding familiar patterns of dementias may be useful to investigators. This section describes how to represent family history using the Associated Persons Medical History (APMH) domain.

2.3.1 Examples for Family History

Example
In this example, subjects were asked about their family history with Alzheimer's disease and MCI in the following manner:

Family History List any first or second degree family members (parents, children, siblings, grandparents, uncles, and aunts) diagnosed with Alzheimer's disease or mild cognitive disorder.				
Relationship to Subject	**Biological Relative? (Y/N)**	**Diagnosed**		**Date of Diagnosis**
		Alzheimer's	**MCI**	
				_ _/_ _/_ _ _ _
				_ _/_ _/_ _ _ _
				_ _/_ _/_ _ _ _
				_ _/_ _/_ _ _ _

MHTERM has been populated with Alzheimer's disease or Mild Cognitive Impairment and MHCAT with PRIMARY DIAGNOSIS. The values in SREL have been populated with terms from the "Relationship to Subject" codelist in the SDTM controlled terminology document.

Rows 1 - 4: Display the Medical History information for relatives of subject AD01-101.
Rows 5 - 7: Display the Medical History information for relatives of subject AD01-102.

apmh.xpt

Row	STUDYID	DOMAIN	APID	MHSEQ	RSUBJID	SREL	MHCAT	MHTERM	MHDTC	MHSTDTC
1	ABC123	APMH	AP_01	1	AD01-101	MOTHER, BIOLOGICAL	PRIMARY DIAGNOSIS	Alzheimer's Disease	2011-03-15	1975-05
2	ABC123	APMH	AP_02	1	AD01-101	SISTER, BIOLOGICAL	PRIMARY DIAGNOSIS	Mild Cognitive Impairment	2011-03-15	2000-07-15
3	ABC123	APMH	AP_03	1	AD01-101	BROTHER, BIOLOGICAL	PRIMARY DIAGNOSIS	Alzheimer's Disease	2011-03-15	1995-07
4	ABC123	APMH	AP_04	1	AD01-101	GRANDMOTHER, BIOLOGICAL	PRIMARY DIAGNOSIS	Alzheimer's Disease	2011-03-15	1940-07
5	ABC123	APMH	AP_05	1	AD01-102	MOTHER, BIOLOGICAL	PRIMARY DIAGNOSIS	Alzheimer's Disease	2011-03-19	1987-01
6	ABC123	APMH	AP_06	1	AD01-102	FATHER BIOLOGICAL	PRIMARY DIAGNOSIS	Mild Cognitive Impairment	2011-03-19	1979-03
7	ABC123	APMH	AP_07	1	AD01-102	GRANDFATHER, BIOLOGICAL	PRIMARY DIAGNOSIS	Alzheimer's Disease	2011-03-19	1945-03

3 Disease Assessments

3.1 Cerebrospinal Fluid Sampling

Cerebrospinal Fluid (CSF) sampling has been identified as an important source of biomarkers as covariates in Alzheimer's disease. The primary biomarkers of interest obtained from these samples include β-amyloid 1-42 (commonly known as Aβ or "Abeta"), tau protein (total), and the phosphorylated form of tau protein. These three markers have been implicated in the pathogenesis of Alzheimer's and Alzheimer's-related dementias. Characterizing a subject's Abeta and tau burden can therefore contribute to the understanding of his or her disease. In this section, we cover how to represent the data generated throughout the process, from sample collection via lumbar puncture (spinal tap), through the aliquoting and storage of the sample, ending with the generation of the results.

It is important to note that several factors of the sample collection methods may affect the results, and therefore should be tightly controlled and recorded. For example, it is possible for Abeta and tau proteins in a CSF sample to react with the material of the sample collection tube, producing differing results in the downstream processing of the same sample if the sample were to be split into two tubes of differing composition. The CSF sample processing examples shown in this guide address this issue from a data standards perspective that shows how to represent the details of these processes in SDTM so that analysts have the necessary information available to interpret the results.

The following concept map shows the provenance of a CSF biomarker value (represented in the LB domain), starting with the generation of the CSF sample via lumbar puncture, then following through the various events the sample may go through (aliquoting, freezing, thawing) in the process of generating actual endpoints. Each step along the process shows at a high-level the different types of data generated that should be represented.

Dashed lines point to the corresponding SDTM domains (or groups of domains, such as those for devices) in which the data generated are submitted. Refer to the individual domains referenced in the concept map below for a detailed view of how the datasets should be populated using more detailed, realistic example values showing β-amyloid, tau, and phosphorylated tau as endpoints. Dashed lines lead to the corresponding SDTM domains in which the data generated are represented. The red lines indicate steps in the process that may occur in differing order: once collected, samples may be processed immediately, or frozen for processing later.

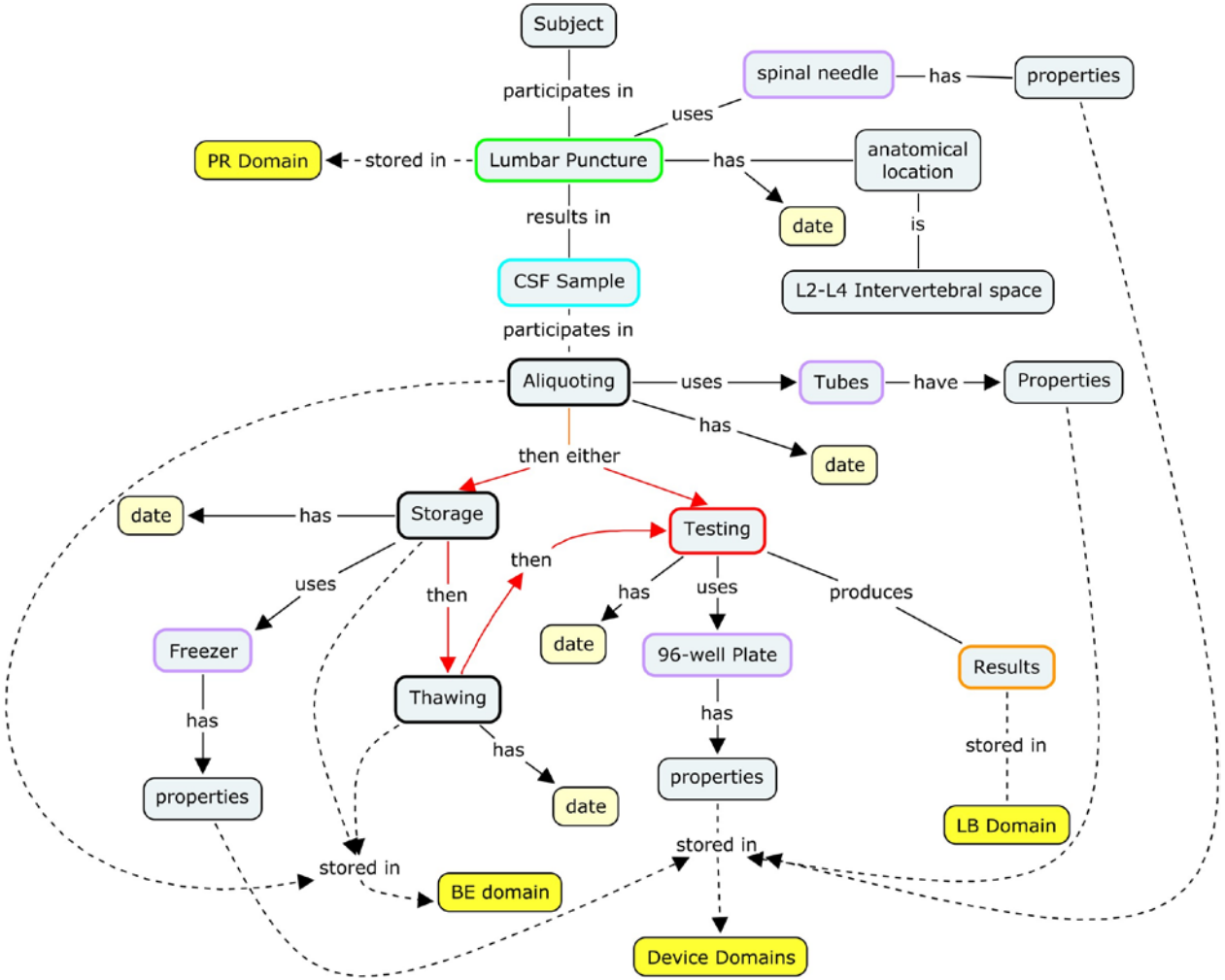


Figure 3.1: CSF Sampling and Processing

3.1.1 Relating Records in CSF Sampling

CSF sampling begins with a lumbar puncture, represented in the Procedures (PR) domain. The origin of the CSF sample in the Biospecimen Events (BE) domain is linked to the lumbar puncture via the REFID variable: PRREFID=BEREFID when BECAT=COLLECTION.

From this point, the CSF sample is usually aliquoted. Each aliquot is assigned a PARENT value that corresponds to the REFID value of the collected sample, thus identifying the aliquot as a "child" of that sample. Additionally, each aliquot would be assigned a unique REFID value that *may* be based on the REFID of the parent. For example, if the PRREFID/BEREFID for the sample collection is 100, then BEREFID for the first aliquoting event could be 100.1, and so on. From this point on, all other events and findings using the sample aliquot will be assigned a REFID value in those subsequent domains corresponding to the unique REFID for the aliquot used. In this way, the REFID values can be viewed as a unique sample ID across all observations recorded on that sample in subsequent domains.

The SPDEVID variable is used to identify devices associated with events or findings in other domains. In the PR domain record for Lumbar Puncture, it identifies the spinal needle used to draw the sample. In the biospecimen domains (BE/BS), it identifies the tube lots used to collect/store the sample and aliquots; it may also be used to represent a separate device identifier for a freezer used for storage when freezing is represented in BE. SPDEVID in the LB domain identifies the 96-well plate used to obtain the ultimate endpoint represented in that domain. All SPDEVID values also match their respective device identifiers and properties represented in the device domains.

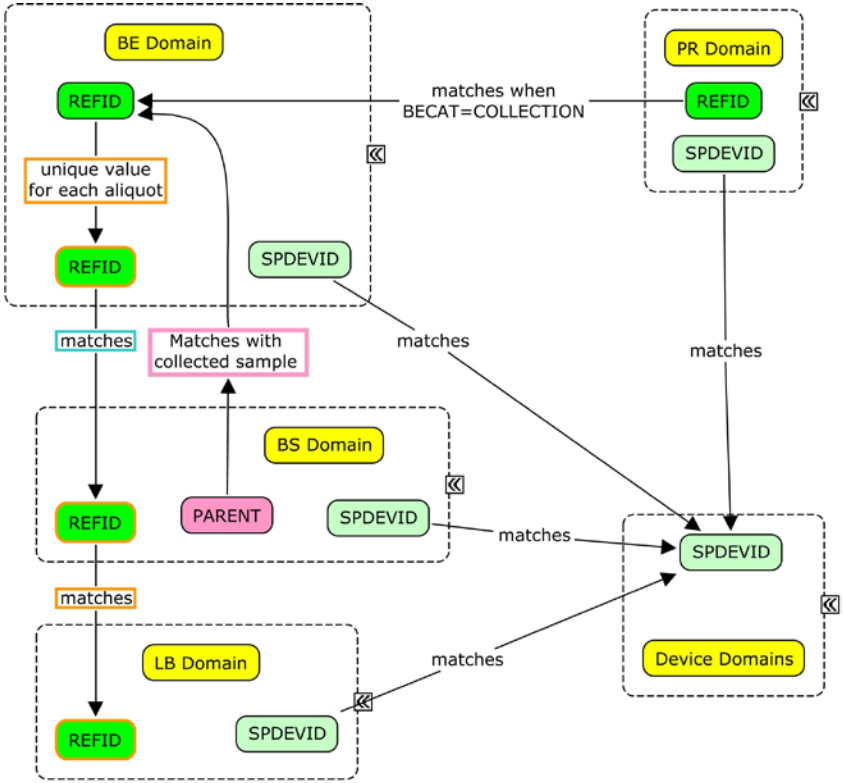


Figure 3.1.1: Relating Records in CSF Sampling

3.1.2 Examples for CSF Sampling

The example below uses domains and variables that are not final at the time of the publication of this document. For additional assumptions, see Appendix C.

Example
This example is organized to reflect the stages of CSF sampling and processing, from collection to results, followed by ancillary device information and dataset relationships.

Because a lumbar puncture may result in adverse events, it is represented in the Procedures (PR) domain**.**

The location of the procedure is noted in PRLOC. The start and end time are noted also. PRREFID will match BEREFID when BECAT=COLLECTION. SPDEVID identifies the spinal needle used to extract the sample. Fasting status (--FAST) is a variable expected to be added to Interventions in SDTM 1.4, and is of special interest for AD studies.

Row 1: The subject was fasting at the time as indicated by PRFAST=Y.

pr.xpt

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	PRSEQ	PRREFID	PRTRT	PRLOC	PRFAST	PRSTDTC	PRENDTC
1	ABC123	PR	AD01-101	23290	1	100	Lumbar Puncture	L3-L4 INTERVERTEBRAL SPACE	Y	2013-05-22T08:15:00	2013-05-22T08:50:00

Because the lumbar puncture is results in a specimen collection, it is also represented in the Biospecimen Event (BE) domain, along with similar information about biospecimen-related events, such as aliquoting or freezing.

For Alzheimer's disease studies, each sample needs a separate identifier to link it to further actions or characteristics of the sample. Therefore, each aliquoting event is assigned a unique BEREFID value that can be traced to the BEREFID assigned for the collected "parent" sample. BEREFID is used to connect the BE and BS (BSREFID) domains, as well as to any results in the LB domain (LBREFID) obtained from the sample.

In this example, one aliquot was thawed and returned to the freezer later. That particular aliquot can be traced through to the LB domain where biomarker test results are represented. For ease of reviewing the example, not all of the detailed rows are shown. These rows have "…" in them. For example, what would have been Rows 6-19 are represented by one row and use the "…" notation. In these rows, only the values for BESEQ and BEREFID change with regard to Row 5.

Row 1: Shows the origin of the specimen (BETERM when BECAT=COLLECTION). Note that BEREFID=100, which matches the PRREFID in PR.xpt above.

Rows 2-20: Show that the sample was aliquoted. Each separate aliquot is assigned a unique BEREFID. In this case, BEREFID is an incremented decimal value based on the original sample using BEREFID when BECAT=COLLECTION as a base number. This is not an explicit requirement, but makes tracking the samples easier. The definitive link between parent-child samples is defined by the PARENT variable shown in the BS domain. SPDEVID refers to the tube storing the aliquot.

Rows 21-39: Show that each aliquot was frozen. BEREFID identifies an individual aliquot. SPDEVID refers to the freezer storing the aliquot.

Rows 40-41: Show an aliquot (BEREFID) that was thawed for analysis then the remainder was re-frozen. Note that the start and end times (BESTDTC, BEENDTC) of the thawing event define the amount of time the sample spent out of the freezer.

be.xpt

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	BESEQ	BEREFID	BETERM	BEDECOD	BECAT	BEDTC	BESTDTC	BEENDTC
1	ABC123	BE	AD01-101	23290	1	100	Lumbar Puncture	LUMBAR PUNCTURE	COLLECTION	2013-05-22	2013-05-22T08:15:00	2013-05-22T08:50:00
2	ABC123	BE	AD01-101	9330	2	100.1	Aliquoted	ALIQUOTED	PREP	2013-05-22	2013-05-22T09:00:00	
3	ABC123	BE	AD01-101	9330	3	100.2	Aliquoted	ALIQUOTED	PREP	2013-05-22	2013-05-22T09:00:00	
4	ABC123	BE	AD01-101	9330	4	100.3	Aliquoted	ALIQUOTED	PREP	2013-05-22	2013-05-22T09:00:00	
5	ABC123	BE	AD01-101	9330	5	100.4	Aliquoted	ALIQUOTED	PREP	2013-05-22	2013-05-22T09:00:00	
…	…	…	…	…	…	…	…	…	…	…	…	…
20	ABC123	BE	AD01-101	9330	20	100.19	Aliquoted	ALIQUOTED	PREP	2013-05-22	2013-05-22T09:00:00	
21	ABC123	BE	AD01-101	31	21	100.1	Frozen	FROZEN	PREP	2013-05-22	2013-05-22T09:11:00	
22	ABC123	BE	AD01-101	31	22	100.2	Frozen	FROZEN	PREP	2013-05-22	2013-05-22T09:11:00	
23	ABC123	BE	AD01-101	31	23	100.3	Frozen	FROZEN	PREP	2013-05-22	2013-05-22T09:11:00	
24	ABC123	BE	AD01-101	31	24	100.4	Frozen	FROZEN	PREP	2013-05-22	2013-05-22T09:11:00	
…	…	…	…	…	…	…	…	…	…	…	…	…
39	ABC123	BE	AD01-101	31	39	100.19	Frozen	FROZEN	PREP	2013-05-22	2013-05-22T09:11:00	
40	ABC123	BE	AD01-101	31	40	100.2	Thawed	THAWED	PREP	2013-05-23	2013-05-23T08:00:00	2013-05-23T08:18:00
41	ABC123	BE	AD01-101	31	41	100.2	Frozen	FROZEN	PREP	2013-05-23	2013-05-23T08:18:00	

Findings data captured during specimen collection, preparation, and handling are represented in the Biospecimen (BS) domain. Note that BS is a draft domain, and the mechanism by which records are related may change.

For ease of use in review, not all of the detailed rows are shown. These rows have "…" in them. For example, Rows 4-19 are represented by one row and use the "…" notation. In these rows, only the values for BSSEQ and BSREFID change with regards to Row 3.

Row 1: Shows the total volume of cerebrospinal fluid collected during the lumbar puncture by using the same values for BSREFID and BEREFID. SPDEVID refers to the lot of tubes used to store the sample. This is the parent (collected) sample from which further aliquots were generated, so BSSPCLVL=1

Rows 2-20: Show the volume of each aliquot created. SPDEVID refers to the lot of tubes used to store the aliquot. BSPARENT for these records is set equal to the BSREFID of the sample shown in Row 1, indicating that the parent sample from which these aliquots were obtained is the 20-mL sample shown in that row. Additionally BSSPCLVL=2 for these rows since these samples are second generation samples generated from the parent sample whose specimen level is 1.

Row 21: Shows the remaining volume of aliquot 2 (REFID=100.2) that was thawed, partially used for analysis, and then re-frozen.

bs.xpt

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	BSSEQ	BSREFID	BSTESTCD	BSTEST	BSCAT	BSORRES	BSORRESU	BSSTRESC
1	ABC123	BS	AD01-101	8842	1	100	VOLUME	Volume	SPECIMEN MEASUREMENT	20	mL	20
2	ABC123	BS	AD01-101	9330	2	100.1	VOLUME	Volume	SPECIMEN MEASUREMENT	2	mL	2
3	ABC123	BS	AD01-101	9330	3	100.2	VOLUME	Volume	SPECIMEN MEASUREMENT	1	mL	1
…	…	…	…	…	…	…	…	…	…	…	…	…
20	ABC123	BS	AD01-101	9330	20	100.19	VOLUME	Volume	SPECIMEN MEASUREMENT	1	mL	1
21	ABC123	BS	AD01-101	9330	21	100.2	VOLUME	Volume	SPECIMEN MEASUREMENT	0.5	mL	0.5

Row	BSSTRESN	BSSTRESU	BSSPEC	BSPARENT	BSSPCLVL	BSDTC
1 (cont)	20	mL	CEREBROSPINAL FLUID		1	2013-05-22
2 (cont)	2	mL	CEREBROSPINAL FLUID	100	2	2013-05-22
3 (cont)	1	mL	CEREBROSPINAL FLUID	100	2	2013-05-22
…	…	…	…	…	…	…
20 (cont)	1	mL	CEREBROSPINAL FLUID	100	2	2013-05-22
21 (cont)	0.5	mL	CEREBROSPINAL FLUID	100	2	2013-05-23

Results from laboratory tests performed on aliquoted samples are represented in the LB domain.

Rows 1-3: Show results for three tests. These results are the second Aliquot of sample 100, as indicated by the value 100.2 in LBREFID. SPDEVID indicates that these tests were done with device 14690. The DI domain has the information that this device is a triplex plate (i.e., a plate containing probes for each of these three biomarkers).

lb.xpt

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	LBSEQ	LBREFID	LBTESTCD	LBTEST	LBCAT	LBORRES	LBORRESU	LBORNRLO	LBORNRHI
1	ABC123	LB	AD01-101	14690	1	100.2	AMYLB42	Amyloid Beta 42	BIOMARKER	400	pg/mL	450	1000
2	ABC123	LB	AD01-101	14690	2	100.2	TPROT	Tau Protein	BIOMARKER	800	pg/mL	125	400
3	ABC123	LB	AD01-101	14690	3	100.2	TPROTP	Phosphorylated Tau Protein	BIOMARKER	120	pg/mL	1	70

Row	LBSTRESC	LBSTRESN	LBSTRESU	LB SPEC	LBBLFL	VISITNUM	LBDTC
1 (cont)	400	400	pg/mL	CSF	Y	1	2013-05-22
2 (cont)	800	800	pg/mL	CSF	Y	1	2013-05-22
3 (cont)	120	120	pg/mL	CSF	Y	1	2013-05-22

Data about the devices used throughout are represented in the Device Identifier (DI) domain.

Rows 1-3: Describe the spinal needle used to extract CSF, the manufacturer and the lot number of the needle.
Rows 4-6: Describe the tube lot used to collect the entire volume of CSF, and the manufacturer of the tubes.
Row 7: Describes the freezer in which the aliquots were stored. This is important for tracking issues related to potential device failures that may affect results.
Rows 8-10: Describe the tube lot used to store the aliquots created from the entire volume sample, and the manufacturer of the lot. All aliquots used the same tube lot.
Rows 11-13: Describe the 96-well plate used to analyze a CSF sample aliquot.

di.xpt

Row	STUDYID	DOMAIN	SPDEVID	DISEQ	DIPARMCD	DIPARM	DIVAL
1	ABC123	DI	23290	1	DEVTYPE	Device Type	Atraumatic Spinal Needle
2	ABC123	DI	23290	2	MANUF	Manufacturer	Acme
3	ABC123	DI	23290	3	LOT	Lot Number	MRC1028
4	ABC123	DI	8842	1	DEVTYPE	Device Type	Tube
5	ABC123	DI	8842	2	MANUF	Manufacturer	Acme
6	ABC123	DI	8842	3	LOT	Lot Number	1200A
7	ABC123	DI	31	1	DEVTYPE	Device Type	Freezer

Row	STUDYID	DOMAIN	SPDEVID	DISEQ	DIPARMCD	DIPARM	DIVAL
8	ABC123	DI	9330	1	DEVTYPE	Device Type	Tube
9	ABC123	DI	9330	2	MANUF	Manufacturer	Acme
10	ABC123	DI	9330	3	LOT	Lot Number	6500A
11	ABC123	DI	14690	1	DEVTYPE	Device Type	Plate
12	ABC123	DI	14690	2	MANUF	Manufacturer	Advantage Biosciences
13	ABC123	DI	14690	3	LOT	Lot Number	1q02192012

The fixed properties of a device identified in DI are represented in the Device Properties (DO) domain.

Row 1: Shows the size of the needle used in the lumbar puncture.
Row 2: Shows the composition of the tube used to store the aliquots.
Row 3: Shows the degree of multiplexing for the 96-well plate.
Row 4: Shows the temperature of the freezer listed in the DI domain. Note that by storing temperature info in the DO domain, it is indicated that this freezer is a dedicated -80 degrees C freezer. If the freezer had been adjustable, the temperature setting would be listed in the DU domain as per the assumptions of these two domains.

do.xpt

Row	STUDYID	DOMAIN	SPDEVID	DOSEQ	DOTESTCD	DOTEST	DOORRES	DOORRESU
1	ABC123	DO	23290	1	SIZE	Size	20	GAUGE
2	ABC123	DO	9330	1	CMPSTN	Composition	Polypropylene	
3	ABC123	DO	14690	1	MULTPLEX	Degree of Multiplexing	Triplex	
4	ABC123	DO	31	1	TEMP	Temperature	-80	C

The RELREC table below shows a number of dataset relationships:

1. How an individual biospecimen (aliquot) event is related to multiple laboratory results records.
2. How an individual CSF sample obtained from the biospecimen collection is related to the lumbar puncture procedure is related to biospecimen collection event for the sample.
3. How the multiple biospecimen events in the BE domain are related to the multiple biospecimen properties in the BS domain.

Rows 1-2: Indicates the link between the biospecimen (aliquot) and the lab result. A unique BEREFID for each aliquot is linked with the same LBREFID value to obtain all the LB results related to the individual aliquot.
Rows 3-4: Show how to relate records between the procedure and the collection of the sample. This describes that the lumbar procedure with PRREFID=100 is linked to the BE collection event with BEREFID=100.
Rows 5-6: Shows the relationship between biospecimen events and biospecimen properties. Note that in the examples above, most records show a one to one relationship (e.g., a volume measurement in BS for each event in BE). However, this relationship can be, and may often be, a many to many relationship. This describes the multiple biospecimen events related to the multiple biospecimen tests that are represented by the RELREC MANY to MANY relationships between the BEREFID and the BSREFID. This relationship is unusual and challenging to manage in a join/merge and only represents the concept of this relationship. To uniquely identify all the BSTESTS related to an individual BETERM, select an individual BEREFID to join/merge with the BSREFID to revise this to a ONE to MANY relationship to obtain all related BSTEST results.

relrec.xpt

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID
1	ABC123	BE		BEREFID		ONE	1
2	ABC123	LB		LBREFID		MANY	1
3	ABC123	PR		PRREFID		ONE	2
4	ABC123	BE		BEREFID		ONE	2
5	ABC123	BE		BEREFID		MANY	3
6	ABC123	BS		BSREFID		MANY	3

3.2 Imaging

Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) or Positron Emission Tomography /Computerized Tomography (PET/CT) are valuable tools for assessing imaging biomarkers in AD. The following subsections describe these procedures at a high level, with an emphasis on the data that are generated, and how they should be represented in SDTM-based datasets.

3.2.1 Magnetic Resonance Imaging

In AD, MRI is primarily used for obtaining volumetric brain measurements including hippocampus volume and total brain volume, as evidence suggests these biomarkers correlate with disease progression. In a brain MRI procedure, the subject lies down and is entered into the MRI scanner. Exogenous contrast-enhancing agents are typically not used in AD. Once the subject is in the scanner, the procedure begins by collecting a series of image slices across the region of interest (in this case, the head). Multiple sets of image slices—commonly referred to as scans—are typically collected across the entire volume of the head, with each new set (scan) potentially using different properties than the scan before. This is an important fact to be aware of since the changing properties between each scan may affect how that image series can be used or interpreted, and is usually controlled by protocol and/or recorded for later reference by analysts. To reiterate: a single MRI procedure (as would be represented in the Procedures (PR) domain) results in multiple scans, each differentiated by the set of properties that were in place for that scan. These changeable settings that may vary from one scan to the next are represented in the Device In-Use (DU) domain. Examples include, but are not limited to, properties such as image weighting, matrix size, slice thickness, repetition time (TR), echo time (TE), inversion time (TI), and anatomical plane of image acquisition. These parameters are usually represented in standardized reports such as the DICOM header.

Additionally, MRI scanners have inherent onboard properties which do not change, including properties such as coil strength. These unchanging properties are represented in the Device Properties (DO) domain. Finally, information such as make, model, and serial number would be represented in the Device Identifiers (DI) domain.

In this section, we show examples of how to represent these types of data in SDTM-based datasets and how to relate them to each other and to the volumetric findings represented in the Morphology (MO) domain. The examples below are only intended to illustrate the data modeling of these various properties and findings and are not meant to serve as strict requirements nor as an all-inclusive list of the properties and findings which should be provided in a regulatory submission. Sponsors should refer to their regulatory agencies/review divisions for guidance on which data should be recorded and submitted as appropriate for their individual studies.

The following concept map shows the provenance of a brain morphological measurement (such as brain volume or hippocampus volume) starting with the generation of the images via MRI and following through to the derivation of the endpoint. Each step along the process shows at a high level the different types of data that are generated and how they should be represented.

For ease of reading, only the salient concepts are shown. Refer to the individual domain examples referenced in the concept map below for a detailed view of how the datasets should be populated.

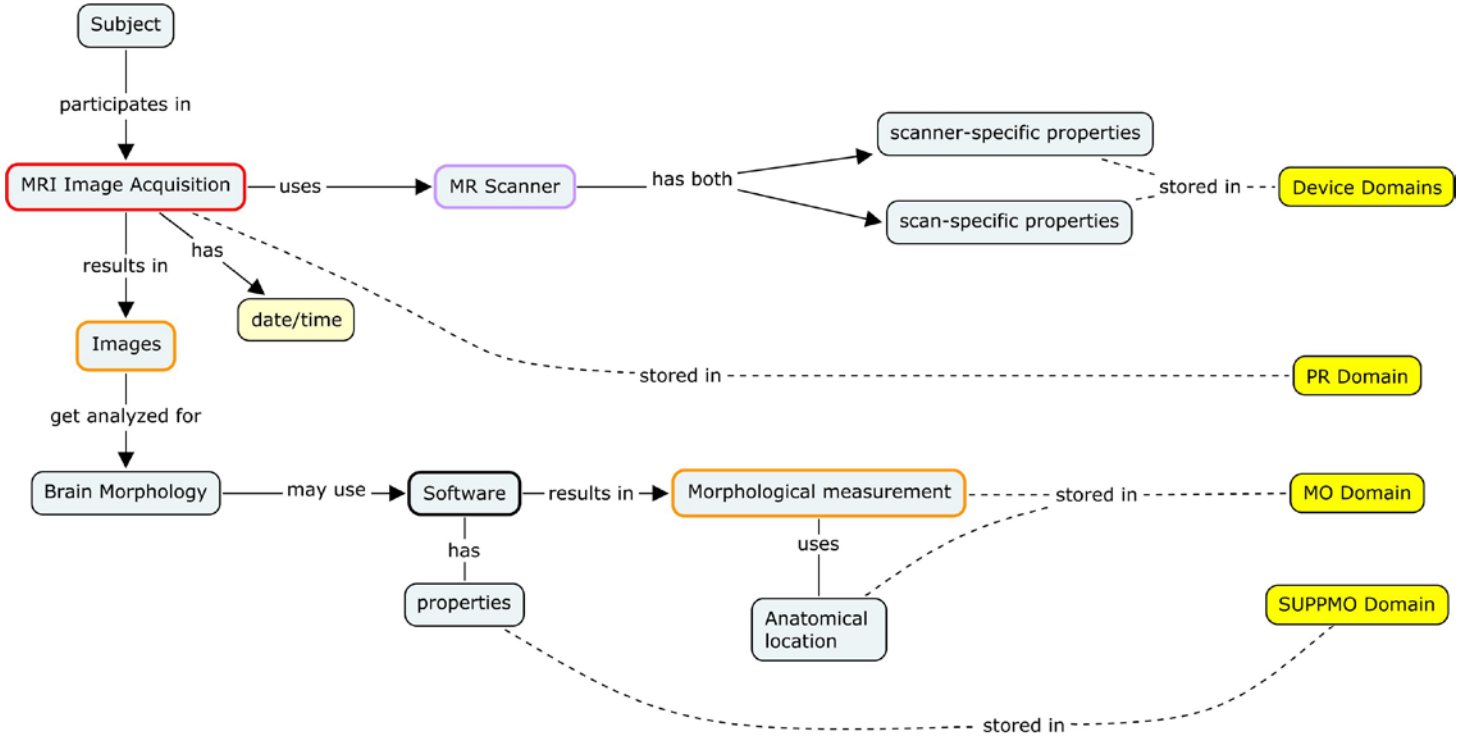


Figure 3.2.1: MRI Imaging

3.2.1.1 Relating Records in MRI Imaging

The derivation of morphological endpoints in MRI imaging begins with the imaging procedure, represented in the PR domain. During this procedure, multiple scans may be conducted. Each scan is identified by a unique set of properties that apply to the scan, and those properties are grouped in the Device In-Use (DU) domain by DUREFID. The link between the findings and the individual scan from which they are derived is maintained by having MOREFID = DUREFID. This allows for linking the findings to a unique scan and its associated parameters within the procedure. A single imaging procedure will have a unique PRREFID, but each scan conducted during that procedure must be differentiated. For this reason, PRREFID cannot be used to link the procedure directly to the findings (brain

volume, etc.) in the MO domain via MOREFID. This direct link between the procedure and any associated findings in MO is maintained by having PRLNKID = MOLNKID. SPDEVID in all cases below refers to the unique identifier for the MRI scanner.

Note that in the use cases examined for Alzheimer's, exogenous contrast reagent administration was not a factor in MRI. If a contrast agent were used, it would be placed in the AG domain according to the example shown in the PET imaging example described in Section 3.2.2.

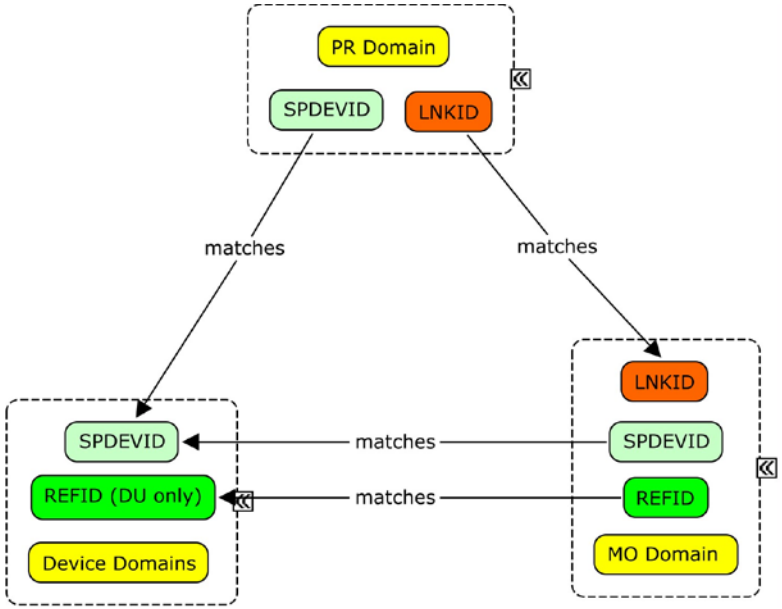


Figure 3.2.1.1: Relating Records in MRI Imaging

3.2.1.2 Examples for MRI Imaging

The example below uses domains and variables which are not final at the time of the publication of this document. For additional assumptions, see Appendix C.

Example
The example below illustrates the modeling of the various pieces of information collected and generated (as described above) from the MRI procedure through the determination of the volumetric results, using the PR, MO, DU, DI, and DO domains. As shown in Figure 3.2.1.1, both --LNKID and --REFID have been used to allow MO to be related to two disparate datasets.

The occurrence of an MRI procedure is represented in the PR domain.

Row 1: Shows an MRI of the head for a subject. PRLNKID links the procedure to the findings in MO via MOLNKID. Note that the subject was fasting during this procedure.

pr.xpt

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	PRSEQ	PRLNKID	PRTRT	PRLOC	PRFAST	PRSTDTC
1	ABC123	PR	AD01-101	16	2	02	MRI	HEAD	Y	2012-05-22T12:30:00

Volumetric endpoints derived from scans during an MRI procedure are represented in the Morphology (MO) domain.

In this example, MOMETHOD indicates that contrast enhancement was not used. If contrast enhancement had been used, MOMETHOD should be set to "CONTRAST ENHANCED MRI" in accordance with controlled terminology. MODTC matches the date of the MRI procedure from which these results were obtained. MOLNKID matches the PRLNKID for the MRI procedure.

Rows 1-3: Show measurements for hippocampal volume taken during the MRI. A measurement for left, right, and both sides is included (MOLAT). The bilateral measure is the sum of the left and right. MOREFID relates the records back to the properties of the originating scan in the DU domain (via DUREFID).

Row 4: Shows a cerebral cortex thickness measurement from the same scan shown in Rows 1-3, as indicated by the same value of MOREFID.

Row 5: Shows a total brain volume measurement from the same patient, but a different scan (as indicated by a different MOREFID). The properties of this scan would also be linked via DUREFID in the DU domain.

mo.xpt

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	MOSEQ	MOLNKID	MOREFID	MOTESTCD	MOTEST	MOORRES	MOORRESU	MOSTRESC	MOSTRESN	MOSTRESU
1	ABC123	MO	AD01-101	16	1	02	1234	VOLUME	Volume	1.8	mL	1.8	1.8	mL
2	ABC123	MO	AD01-101	16	2	02	1234	VOLUME	Volume	1.9	mL	1.9	1.9	mL
3	ABC123	MO	AD01-101	16	3	02	1234	VOLUME	Volume	3.7	mL	3.7	3.7	mL
4	ABC123	MO	AD01-101	16	4	02	1234	THICK	Thickness	3	mm	3	3	mm
5	ABC123	MO	AD01-101	16	5	02	1235	VOLUME	Volume	864	mL	864	864	mL

Row	MOLOC	MOLAT	MOMETHOD	MOANMETH	MODTC
1 (cont)	HIPPOCAMPUS	RIGHT	MRI WITHOUT CONTRAST	SOFTWARE ANAYLSIS	2012-05-22
2 (cont)	HIPPOCAMPUS	LEFT	MRI WITHOUT CONTRAST	SOFTWARE ANAYLSIS	2012-05-22
3 (cont)	HIPPOCAMPUS	BILATERAL	MRI WITHOUT CONTRAST	SOFTWARE ANAYLSIS	2012-05-22
4 (cont)	CEREBRAL CORTEX		MRI WITHOUT CONTRAST	SOFTWARE ANAYLSIS	2012-05-22
5 (cont)	BRAIN		MRI WITHOUT CONTRAST	SOFTWARE ANAYLSIS	2012-05-22

Because two devices are used in the MRI procedure, the MRI device is represented in SPDEVID, while the analysis software used is represented using a Supplemental Qualifiers dataset. This method is under review by the CDISC standards review council (SRC) and should not be considered final or definitive. Future decisions by SRC may require recording the analysis software as a device as well.

suppmo.xpt

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
1	ABC123	MO	AD01-101	MOSEQ	1	SFTWR	Analysis Software	Neuroquant
2	ABC123	MO	AD01-101	MOSEQ	1	SFTWRVER	Software Version	12.1
3	ABC123	MO	AD01-101	MOSEQ	2	SFTWRE	Analysis Software	Neuroquant
4	ABC123	MO	AD01-101	MOSEQ	2	SFTWRVER	Software Version	12.1
5	ABC123	MO	AD01-101	MOSEQ	3	SFTWR	Analysis Software	Neuroquant

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
6	ABC123	MO	AD01-101	MOSEQ	3	SFTWRVER	Software Version	12.1
7	ABC123	MO	AD01-101	MOSEQ	4	SFTWR	Analysis Software	Neuroquant
8	ABC123	MO	AD01-101	MOSEQ	4	SFTWRVER	Software Version	12.1
9	ABC123	MO	AD01-101	MOSEQ	5	SFTWR	Analysis Software	Freesurfer
10	ABC123	MO	AD01-101	MOSEQ	5	SFTWRVER	Software Version	4.0.2

The MRI imaging device used, the manufacturer, and the model are represented in the DI domain. Because SPDEVID is included as a variable in the other datasets, RELREC is not needed to relate DI to other domains.

di.xpt

Row	STUDYID	DOMAIN	SPDEVID	DISEQ	DIPARMCD	DIPARM	DIVAL
1	ABC123	DI	16	1	DEVTYPE	Device Type	MRI
2	ABC123	DI	16	2	MANUF	Manufacturer	GE
3	ABC123	DI	16	3	MODEL	Model	SIGNA

Changeable properties and parameters for devices identified in DI are represented in the Device In-Use (DU) domain.

The example below shows how to represent various properties of the MRI device that may change between scans. DUREFID indicates a unique scan, as defined by a changeable set of properties that were set as indicated for each scan. Note that the properties in this example are meant to serve as *examples only* of properties that could appear in a DICOM header. Sponsors should refer to FDA guidance and speak to their individual review division(s) to determine which parameters should be included in the data submission, and then verify if controlled terminology exists for these parameters. New terminology for the DUTEST and DUTESTCD code lists can be requested through the new term request process at NCI EVS.

Rows 1-16: This particular scan was conducted in the sagittal plane, using the MRI scanner represented by SPDEVID=16.

Rows 17-32: This particular scan was conducted in the transverse plane. This represents the second scan conducted on the same subject during the same imaging procedure that started with the scan depicted in rows 1-16. The only property that has changed for this scan from the previous scan is represented in DUTEST=Anatomical Plane.

du.xpt

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	DUSEQ	DUREFID	DUTESTCD	DUTEST	DUORRES	DUORRESU	…	VISITNUM	DUDTC
1	ABC123	DU	AD01-101	16	1	1234	ANTPLANE	Anatomical Plane	SAGITTAL		…	1	2012-05-22T12:30:00
2	ABC123	DU	AD01-101	16	2	1234	PULSSEQ	Pulse Sequence	SPGR		…	1	2012-05-22T12:30:00
3	ABC123	DU	AD01-101	16	3	1234	STHICK	Slice Thickness	4	mm	…	1	2012-05-22T12:30:00
4	ABC123	DU	AD01-101	16	4	1234	AQMTRXSZ	Image Acquisition Matrix Size	256X256		…	1	2012-05-22T12:30:00
5	ABC123	DU	AD01-101	16	5	1234	WEIGHTNG	Weighting	T1		…	1	2012-05-22T12:30:00
6	ABC123	DU	AD01-101	16	6	1234	INTSPACE	Interslice Spacing	1	mm	…	1	2012-05-22T12:30:00
7	ABC123	DU	AD01-101	16	7	1234	FLDVIEW	Field of View	280X280	mm	…	1	2012-05-22T12:30:00
8	ABC123	DU	AD01-101	16	8	1234	SFTWRVER	Software Version	3.1		…	1	2012-05-22T12:30:00
9	ABC123	DU	AD01-101	16	9	1234	FLIPANGL	Flip Angle	8	deg	…	1	2012-05-22T12:30:00
10	ABC123	DU	AD01-101	16	10	1234	PIXSPCX	Pixel Spacing X	1.25	mm	…	1	2012-05-22T12:30:00
11	ABC123	DU	AD01-101	16	11	1234	PIXSPCY	Pixel Spacing Y	1.25	mm	…	1	2012-05-22T12:30:00
12	ABC123	DU	AD01-101	16	12	1234	INVRTIME	Inversion Time	1000	msec	…	1	2012-05-22T12:30:00
13	ABC123	DU	AD01-101	16	13	1234	ECHOTIME	Echo Time	3.59	msec	…	1	2012-05-22T12:30:00
14	ABC123	DU	AD01-101	16	14	1234	REPTIME	Repetition Time	3000	msec	…	1	2012-05-22T12:30:00

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	DUSEQ	DUREFID	DUTESTCD	DUTEST	DUORRES	DUORRESU	…	VISITNUM	DUDTC
15	ABC123	DU	AD01-101	16	15	1234	NUMSLICE	Number of Slices	125		…	1	2012-05-22T12:30:00
16	ABC123	DU	AD01-101	16	16	1234	IMAQDIM	Image Acquisition Dimensionality	3D		…	1	2012-05-22T12:30:00
17	ABC123	DU	AD01-101	16	17	1235	ANTPLANE	Anatomical Plane	TRANSVERSE		…	1	2012-05-22T12:50:00
18	ABC123	DU	AD01-101	16	18	1235	PULSSEQ	Pulse Sequence	SPGR		…	1	2012-05-22T12:50:00
19	ABC123	DU	AD01-101	16	19	1235	STHICK	Slice Thickness	4	mm	…	1	2012-05-22T12:50:00
20	ABC123	DU	AD01-101	16	20	1235	AQMTRXSZ	Image Acquisition Matrix Size	256X256	Pixels	…	1	2012-05-22T12:50:00
21	ABC123	DU	AD01-101	16	21	1235	WEIGHTNG	Weighting	T1		…	1	2012-05-22T12:50:00
22	ABC123	DU	AD01-101	16	22	1235	INTSPACE	Interslice Spacing	1	mm	…	1	2012-05-22T12:50:00
23	ABC123	DU	AD01-101	16	23	1235	FLDVIEW	Field of View	280X280	mm	…	1	2012-05-22T12:50:00
24	ABC123	DU	AD01-101	16	24	1235	SFTWRVER	Software Version	3.1		…	1	2012-05-22T12:50:00
25	ABC123	DU	AD01-101	16	25	1235	FLIPANGL	Flip Angle	8	deg	…	1	2012-05-22T12:50:00
26	ABC123	DU	AD01-101	16	26	1235	PIXSPCX	Pixel Spacing X	1.25	mm	…	1	2012-05-22T12:50:00
27	ABC123	DU	AD01-101	16	27	1235	PIXSPCY	Pixel Spacing Y	1.25	mm	…	1	2012-05-22T12:50:00
28	ABC123	DU	AD01-101	16	28	1235	INVRTIME	Inversion Time	1000	msec	…	1	2012-05-22T12:50:00
29	ABC123	DU	AD01-101	16	29	1235	ECHOTIME	Echo Time	3.59	msec	…	1	2012-05-22T12:50:00
30	ABC123	DU	AD01-101	16	30	1235	REPTIME	Repetition Time	3000	msec	…	1	2012-05-22T12:50:00
31	ABC123	DU	AD01-101	16	31	1235	NUMSLICE	Number of Slices	125		…	1	2012-05-22T12:50:00
32	ABC123	DU	AD01-101	16	32	1235	IMAQDIM	Image Acquisition Dimensionality	3D		…	1	2012-05-22T12:50:00

Fixed properties for the MRI scanner used are represented in the DO domain.

do.xpt

Row	STUDYID	DOMAIN	SPDEVID	DOSEQ	DOTESTCD	DOTEST	DOORRES	DOORRESU
1	ABC123	DO	16	1	COILTYPE	Coil Type	Head	
2	ABC123	DO	16	2	COILSTR	Coil Strength	1.5	T

The relationship between the MRI scan and the results, and the relationship between the results and the settings of the MRI device, are represented in RELREC.

Rows 1-2: Link the morphology result to the particular procedure performed. This indicates that the PRLNKID variable for an individual procedure is linked to multiple MO records via the same value in the MOLNKID variable.

Rows 3-4: Link the morphology result to the particular scan and properties/settings of the scanner. This describes the multiple MO results that are related to multiple DUTESTs that are represented by the RELREC MANY to MANY relationships between the MOREFID and the DUREFID. This relationship is unusual and challenging to manage in a join/merge and only represents the concept of this relationship. To uniquely identify all the DUTESTS related to an individual MOTEST, first select an individual MOREFID to join/merge with the DUREFID to revise this to a ONE to MANY relationship to obtain all related DUTEST results.

relrec.xpt

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID
1	ABC123	PR		PRLNKID		ONE	5
2	ABC123	MO		MOLNKID		MANY	5
3	ABC123	MO		MOREFID		MANY	4
4	ABC123	DU		DUREFID		MANY	4

3.2.2 Positron Emission Tomography Imaging

Unlike the volumetric MRI imaging described in Section 3.2.1, PET imaging is used to characterize functional processes in the brain. In AD, it is used to quantify amyloid plaques and to assess glucose metabolism. PET imaging requires the intravenous administration of a radiopharmaceutical, or radiolabeled "tracer," to the subject in order to target the region of interest for imaging. Tracers are specific to the processes they are designed to characterize. A tracer designed for amyloid plaque imaging would not work for assessment of glucose metabolism, nor vice versa. These tracers are molecules designed to be taken up by or bind to specific regions of the brain and contain a radioactive element. When the tracer accumulates in the region it is designed to bind, the radioactive signal emitted from this region rises relative to the signal emitted from regions it is not designed to bind. PET scanners are designed to collect the radioactivity emitted and localize this signal to the region where it has accumulated. These tracers require a specific uptake time—the time from when they were administered to the subject to when the imaging procedure can begin—in order to ensure time is given for them to accumulate at their target location in the brain. Consistent uptake times are required to ensure comparability of results across subjects and visits within a given protocol. These tracers are represented in the Procedure Agents (AG) domain. It is important to record the start time of tracer administration, as well as the start time of the imaging procedure (represented, as in the MRI examples, in the PR domain). This ensures analysts can determine if uptake time was comparable across observations.

Once the specified uptake time has passed, the patient is entered into the PET scanner, and the imaging procedure begins. As in MRI, a single PET procedure may result in multiple scans. As in the MRI examples, the set of changeable properties that define a scan are grouped in the DU domain, and linked to the findings in the NV domain, by using the same --REFID value. PET scanners also have unchanging properties and identifiers analogous to MRI scanners, which are represented in the DO and DI domains, respectively, as described in Section 3.2.1.

The ultimate endpoint generated by a PET procedure comes in the form of a Standard Uptake Value Ratio (SUVR). SUVR is a ratio of the signal coming from the brain region of interest (to which the tracer is designed to target) compared to the signal emitted from some reference region in the brain. Reference regions are chosen based on the likelihood that the tracer will not be taken up significantly in that region, and thus they serve as a background signal for normalization. Since SUVR is a surrogate measurement for some functional activity, as opposed to a mere structural (morphological) measurement as in the case of volumetric MRI endpoints, these results are represented in the Nervous System Findings (NV) domain.

The anatomical data provided by PET scanners are often not sufficient to localize the radioactivity they are designed to image to the level of specificity desired by investigators. For this reason, PET scans are often conducted concurrently with CT scans, which allow co-registration of more detailed anatomical data with the SUVR data obtained from the PET portion of the image. This is performed concurrently in a single procedure using a hybrid PET/CT scanner. In the use cases examined for this guide, the CT portion of the image was used solely for the purpose of anatomical co-registration of the PET data from which the endpoints of interest were obtained. Therefore, specific properties of the CT portion of the scans are not represented in the device domains in this guide. However, the use of a hybrid PET/CT scanner should be represented in the device type value of the DI domain, and as the method value in the NV domain.

It is important to realize that SUVR is based on quantifying regions of signal in a series of images. The only things that differentiate one SUVR from the next are the region of interest and the radiolabeled tracer used. It is therefore essential to record the region of interest and represent this in the NVLOC variable. Associated qualifiers of NVLOC (NVLAT, NVDIR) may be used where appropriate. It is also essential to maintain the link between the SUVR value (in the NV domain) and the tracer that produced it (in the AG domain).

In the following examples, we demonstrate how to represent these types of data in SDTM-based datasets, and how to maintain and represent the relationship between the tracer administration, the PET procedure, the individual scan properties, and the SUVR findings represented in the NV domain. The examples below are only intended to illustrate the data modeling of these various properties and findings, and are not meant to serve as strict requirements or an all-

inclusive list of the properties and findings which should be provided in a regulatory submission. Sponsors should refer to their regulatory agencies/review divisions for guidance on which data should be recorded and submitted as appropriate to their individual protocol(s).

The following concept map shows the provenance of a brain physiological or functional measurement (measured as SUVR, represented in the NV domain) of various PET tracers. The process starts with the administration of the radiopharmaceutical, followed by the acquisition of images, and ultimately, the derivation of the endpoint. Each step along the process shows in high-level detail the different types of data generated that should be represented.
Dashed lines lead to the corresponding SDTM domains in which the data generated are represented.

For ease of reading, only the salient concepts are shown. Refer to the individual domain examples referenced in the concept map below for detailed views of how the datasets should be populated.

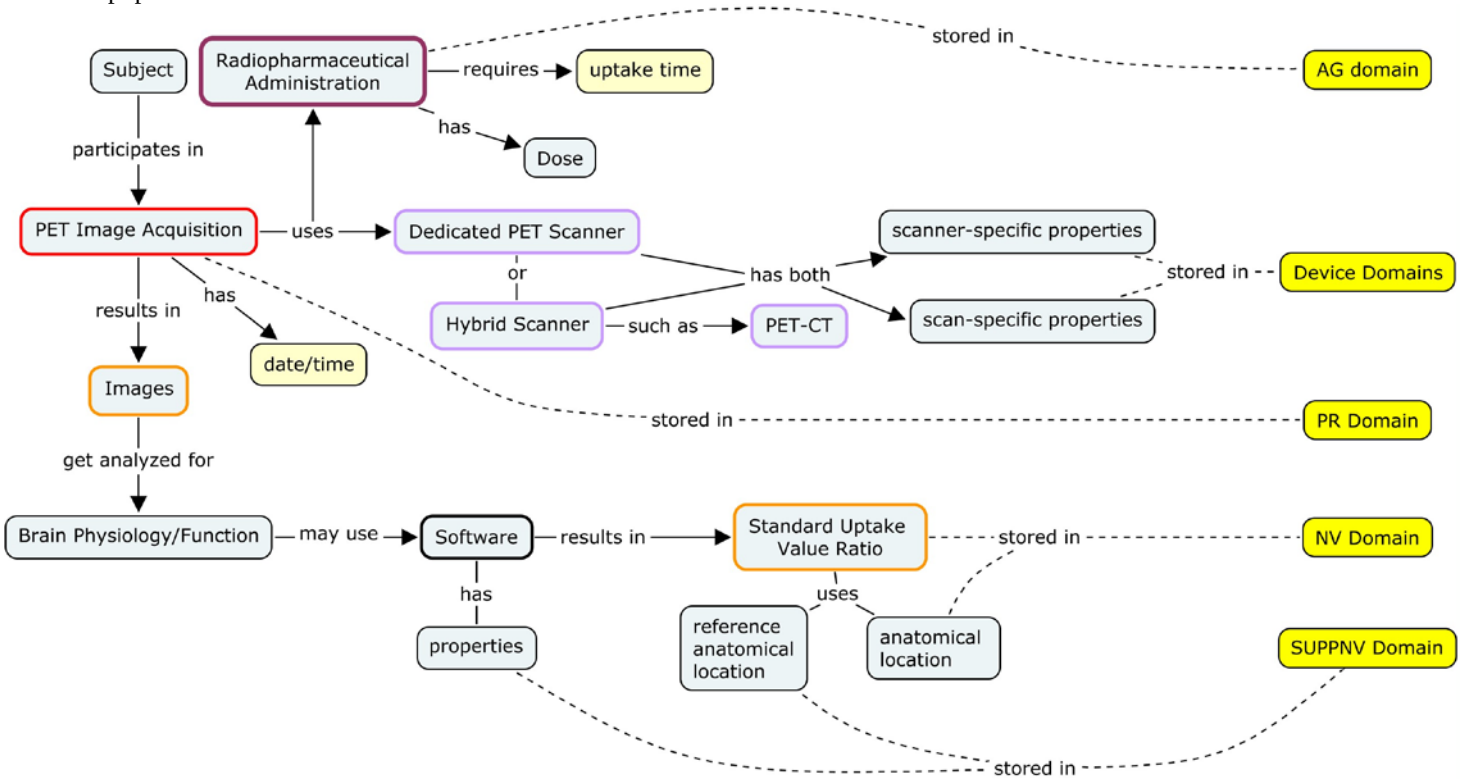


Figure 3.2.2: PET Imaging

3.2.2.1 Relating Records in PET Imaging

Similar to MRI imaging, PET imaging begins with the imaging procedure, represented in the PR domain. The structure and relationships of the data generated are handled exactly as in the MRI example, except that the findings are functional or physiological, not morphological, and are therefore represented in the NV domain. Furthermore, all PET use cases examined involved the administration of a radiolabeled tracer, which is represented in the AG domain. In these cases, the LNKID variable serves as a three-way link between the procedure, the tracer administration, and the results. It is particularly important to maintain the link between AGLNKID and NVLNKID as the tracer information in the AG domain describes the target biomarker being imaged, and on which the SUVR in the NV domain is based. Note that this approach would also work for MRI use cases that involve an exogenous contrast-agent administration. In the PET and PET/CT examples discussed, SPDEVID refers to the PET or PET/CT scanner used.

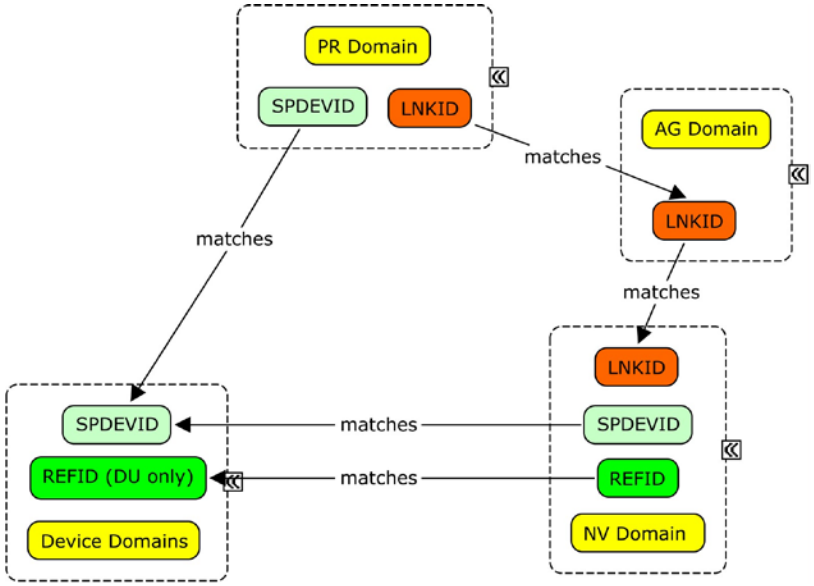


Figure 3.2.2.1: Relating Records in PET Imaging

3.2.2.2 Examples for PET Imaging

The example below uses domains and variables that are not final at the time of the publication of this document. For additional assumptions, see Appendix C.

Example
This example shows the modeling of the various pieces of information collected and generated (as described above) from three separate PET or PET/CT procedures, each for a separate subject. Two procedures (for subjects AD01-101 and AD01-102) showcase amyloid imaging using two different tracers (florbetapir and 11C-PiB), while the third (for subject AD01-103) shows glucose-metabolism imaging using FDG as the tracer. Data include the procedure

records in the PR domain, the administration of the tracer in the AG domain, the SUVR endpoints in the NV domain, and the various device properties in the respective domains. All subjects were fasting at the time of the procedures as indicated by PRFAST=Y. SPDEVID identifies the scanner used in the procedure.

Occurrences of PET or PET/CT procedures are recorded in the PR domain.

Rows 1-2: Show PET/CT head scans for two subjects using the same PET/CT scanner, as defined in SPDEVID.
Row 3: Shows a PET scan of the head for a third subject.

pr.xpt

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	PRSEQ	PRLNKID	PRTRT	PRLOC	PRFAST	PRSTDTC
1	ABC123	PR	AD01-101	22	1	03	PET/CT	HEAD	Y	2012-05-22T09:30:00
2	ABC123	PR	AD01-102	22	1	04	PET/CT	HEAD	Y	2012-05-22T08:00:00
3	ABC123	PR	AD01-103	44	1	05	PET	HEAD	Y	2012-05-22T09:00:00

Data about the administration of procedure agents such as a radiopharmaceutical tracer for PET or PET/CT scans are represented in the Procedure Agents (AG) domain.

Rows 1-3: Show timing and dosage for tracer substances injected during a PET scan. This relationship is shown in the RELREC table. In this example, AGCAT is used to group tracers with the same target.

ag.xpt

Row	STUDYID	DOMAIN	USUBJID	AGSEQ	AGLNKID	AGTRT	AGCAT	AGDOSE	AGDOSEU	AGSTDTC
1	ABC123	AG	AD01-101	1	03	18F-Florbetapir	AMYLOID TRACER	370	MBq	2012-05-22T08:40:00
2	ABC123	AG	AD01-102	1	04	11C-PiB	AMYLOID TRACER	370	MBq	2012-05-22T07:20:00
3	ABC123	AG	AD01-103	1	05	FDG	GLUCOSE TRACER	400	MBq	2012-05-22T08:30:00

Data about the nervous system physiology or metabolism, such as measured by PET imaging, are represented in the Nervous System Findings (NV) domain.

This example shows measures for standard uptake value ratios taken from three PET scans. SPDEVID shows the scanner used. NVLNKID can be used to link back to the imaging procedure record in the PR domain (PRLNKID), as well as to the tracer administration record in the AG domain (AGLNKID). AGLNKID would be used to determine which tracer uptake is being measured (SUVR), and therefore to which biomarker the findings pertain. NVDTC corresponds to the date of the PET or PET/CT procedure from which these results were obtained.

Rows 1-2: Show the Standard Uptake Value Ratio (SUVR) findings based on a PET/CT scan for subject AD01-101.
Rows 3-4: Show the SUVR findings based on a PET/CT scan for subject AD01-102.
Rows 5-6: Show the SUVR findings based on an FDG-PET scan for subject AD AD01-103.

nv.xpt

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	NVSEQ	NVREFID	NVLNKID	NVTESTCD	NVTEST	NVORRES	NVORRESU
1	ABC123	NV	AD01-101	22	1	1236	03	SUVR	Standard Uptake Value Ratio	.95	RATIO

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	NVSEQ	NVREFID	NVLNKID	NVTESTCD	NVTEST	NVORRES	NVORRESU
2	ABC123	NV	AD01-101	22	2	1236	03	SUVR	Standard Uptake Value Ratio	1.17	RATIO
3	ABC123	NV	AD01-102	22	1	1237	04	SUVR	Standard Uptake Value Ratio	1.21	RATIO
4	ABC123	NV	AD01-102	22	2	1237	04	SUVR	Standard Uptake Value Ratio	1.78	RATIO
5	ABC123	NV	AD01-103	44	1	1238	05	SUVR	Standard Uptake Value Ratio	1.52	RATIO
6	ABC123	NV	AD01-103	44	2	1238	05	SUVR	Standard Uptake Value Ratio	1.63	RATIO

Row	NVSTRESC	NVSTRESN	NVSTRRESU	NVLOC	NVDIR	NVMETHOD	NVDTC
1 (cont)	.95	.95	RATIO	PRECUNEUS		PET/CT SCAN	2012-05-22
2 (cont)	1.17	1.17	RATIO	CINGULATE CORTEX	POSTERIOR	PET/CT SCAN	2012-05-22
3 (cont)	1.21	1.21	RATIO	PRECUNEUS		PET/CT SCAN	2012-05-22
4 (cont)	1.78	1.78	RATIO	CINGULATE CORTEX	POSTERIOR	PET/CT SCAN	2012-05-22
5 (cont)	1.52	1.52	RATIO	PRECUNEUS		FDGPET	2012-05-22
6 (cont)	1.63	1.63	RATIO	CINGULATE CORTEX	POSTERIOR	FDGPET	2012-05-22

A Supplemental Qualifiers dataset is needed for additional data elements that are not part of the NV domain.

Rows 1-6: Shows the reference region used for the SUVR tests shown in the NV domain.

suppnv.xpt

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
1	ABC123	NV	AD01-101	NVSEQ	1	REFREG	Reference Region	CEREBELLUM
2	ABC123	NV	AD01-101	NVSEQ	2	REFREG	Reference Region	CEREBELLUM
3	ABC123	NV	AD01-102	NVSEQ	1	REFREG	Reference Region	CEREBELLUM
4	ABC123	NV	AD01-102	NVSEQ	2	REFREG	Reference Region	CEREBELLUM
5	ABC123	NV	AD01-103	NVSEQ	1	REFREG	Reference Region	PONS VAROLII
6	ABC123	NV	AD01-103	NVSEQ	2	REFREG	Reference Region	PONS VAROLII

Data about the devices used throughout are represented in the DI domain.

Rows 1-3: Describe the PET/CT imaging device used in the Amyloid Tracer examples shown in the NV domain.
Rows 4-6: Describe the PET imaging device used in the Glucose Tracer (FDG) example shown in the NV domain.

di.xpt

Row	STUDYID	DOMAIN	SPDEVID	DISEQ	DIPARMCD	DIPARM	DIVAL
1	ABC123	DI	22	1	DEVTYPE	Device Type	PET/CT
2	ABC123	DI	22	2	MANUF	Manufacturer	Siemens
3	ABC123	DI	22	3	MODEL	Model	TRIO
4	ABC123	DI	44	1	DEVTYPE	Device Type	PET
5	ABC123	DI	44	2	MANUF	Manufacturer	Siemens
6	ABC123	DI	44	3	MODEL	Model	INVEON

Changeable properties and parameters of the devices identified in DI are represented in the DU domain. DUREFID is used to identify each unique scan, as defined by a changeable set of properties that were set as indicated for this scan.

Rows 1-13: Show properties of the PET/CT device that were used in a scan of subject AD01-101.
Rows 14-26: Show properties of the PET/CT device that were used in a scan of subject AD01-102.
Rows 27-39: Show properties of the PET device that were used in a scan of subject AD01-103.

du.xpt

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	DUSEQ	DUREFID	DUTESTCD	DUTEST	DUORRES	DUORRESU	…	VISITNUM	DUDTC
1	ABC123	DU	AD01-101	22	1	1236	ANTPLANE	Anatomical Plane	SAGITTAL		…	1	2012-05-22T09:30:00
2	ABC123	DU	AD01-101	22	2	1236	INTSPACE	Interslice Spacing	1	mm	…	1	2012-05-22T09:30:00
3	ABC123	DU	AD01-101	22	3	1236	SFTWRVER	Software Version	5.1		…	1	2012-05-22T09:30:00
4	ABC123	DU	AD01-101	22	4	1236	STHICK	Slice Thickness	5	mm	…	1	2012-05-22T09:30:00
5	ABC123	DU	AD01-101	22	5	1236	PIXSPCX	Pixel Spacing X	2	mm	…	1	2012-05-22T09:30:00
6	ABC123	DU	AD01-101	22	6	1236	PIXSPCY	Pixel Spacing Y	2	mm	…	1	2012-05-22T09:30:00
7	ABC123	DU	AD01-101	22	7	1236	AQMTRXSZ	Image Acquisition Matrix Size	256X256		…	1	2012-05-22T09:30:00
8	ABC123	DU	AD01-101	22	8	1236	FLDVIEW	Field of View	280X280	mm	…	1	2012-05-22T09:30:00
9	ABC123	DU	AD01-101	22	9	1236	NUMSLICE	Number of Slices	125		…	1	2012-05-22T09:30:00
10	ABC123	DU	AD01-101	22	10	1236	ATTCRCT	Attenuation Correction Type	FBP		…	1	2012-05-22T09:30:00
11	ABC123	DU	AD01-101	22	11	1236	DECCORR	Decay Correction	N		…	1	2012-05-22T09:30:00
12	ABC123	DU	AD01-101	22	12	1236	RANDCORR	Randoms Correction	N		…	1	2012-05-22T09:30:00
13	ABC123	DU	AD01-101	22	13	1236	RECONDAT	Reconstruction of Raw Data Type	Iterative		…	1	2012-05-22T09:30:00
14	ABC123	DU	AD01-102	22	1	1237	ANTPLANE	Anatomical Plane	SAGITTAL		…	1	2012-05-22T08:00:00
15	ABC123	DU	AD01-102	22	2	1237	INTSPACE	Interslice Spacing	1	mm	…	1	2012-05-22T08:00:00
16	ABC123	DU	AD01-102	22	3	1237	SFTWRVER	Software Version	5.1		…	1	2012-05-22T08:00:00
17	ABC123	DU	AD01-102	22	4	1237	STHICK	Slice Thickness	5	mm	…	1	2012-05-22T08:00:00
18	ABC123	DU	AD01-102	22	5	1237	PIXSPCX	Pixel Spacing X	2	mm	…	1	2012-05-22T08:00:00
19	ABC123	DU	AD01-102	22	6	1237	PIXSPCY	Pixel Spacing Y	2	mm	…	1	2012-05-22T08:00:00
20	ABC123	DU	AD01-102	22	7	1237	AQMTRXSZ	Image Acquisition Matrix Size	256X256		…	1	2012-05-22T08:00:00
21	ABC123	DU	AD01-102	22	8	1237	FLDVIEW	Field of View	280X280	mm	…	1	2012-05-22T08:00:00
22	ABC123	DU	AD01-102	22	9	1237	NUMSLICE	Number of Slices	125		…	1	2012-05-22T08:00:00
23	ABC123	DU	AD01-102	22	10	1237	ATTCRCT	Attenuation Correction Type	FBP		…	1	2012-05-22T08:00:00
24	ABC123	DU	AD01-102	22	11	1237	DECCORR	Decay Correction	N		…	1	2012-05-22T08:00:00
25	ABC123	DU	AD01-102	22	12	1237	RANDCORR	Randoms Correction	N		…	1	2012-05-22T08:00:00
26	ABC123	DU	AD01-102	22	13	1237	RECONDAT	Reconstruction of Raw Data Type	Backscatter		…	1	2012-05-22T08:00:00
27	ABC123	DU	AD01-103	44	1	1238	ANTPLANE	Anatomical Plane	SAGITTAL		…	1	2012-05-22T09:30:00
28	ABC123	DU	AD01-103	44	2	1238	INTSPACE	Interslice Spacing	1	mm	…	1	2012-05-22T09:30:00
29	ABC123	DU	AD01-103	44	3	1238	SFTWRVER	Software Version	5.1		…	1	2012-05-22T09:30:00
30	ABC123	DU	AD01-103	44	4	1238	STHICK	Slice Thickness	5	mm	…	1	2012-05-22T09:30:00
31	ABC123	DU	AD01-103	44	5	1238	PIXSPCX	Pixel Spacing X	2	mm	…	1	2012-05-22T09:30:00
32	ABC123	DU	AD01-103	44	6	1238	PIXSPCY	Pixel Spacing Y	2	mm	…	1	2012-05-22T09:30:00
33	ABC123	DU	AD01-103	44	7	1238	AQMTRXSZ	Image Acquisition Matrix Size	256X256		…	1	2012-05-22T09:30:00
34	ABC123	DU	AD01-103	44	8	1238	FLDVIEW	Field of View	280X280	mm	…	1	2012-05-22T09:30:00
35	ABC123	DU	AD01-103	44	9	1238	NUMSLICE	Number of Slices	125		…	1	2012-05-22T09:30:00
36	ABC123	DU	AD01-103	44	10	1238	ATTCRCT	Attenuation Correction Type	FBP		…	1	2012-05-22T09:30:00
37	ABC123	DU	AD01-103	44	11	1238	DECCORR	Decay Correction	N		…	1	2012-05-22T09:30:00
38	ABC123	DU	AD01-103	44	12	1238	RANDCORR	Randoms Correction	N		…	1	2012-05-22T09:30:00
39	ABC123	DU	AD01-103	44	13	1238	RECONDAT	Reconstruction of Raw Data Type	Backscatter		…	1	2012-05-22T09:30:00

Fixed properties of devices identified in DI are represented in the DO domain.

Row 1: Shows properties of the PET/CT scanner used.
Row 2: Shows properties of the PET scanner used.
Row 3: Shows the composition of the tube used to collect the CSF sample from the lumbar puncture in the PR domain.

do.xpt

Row	STUDYID	DOMAIN	SPDEVID	DOSEQ	DOTESTCD	DOTEST	DOORRES	DOORRESU
1	ABC123	DO	22	1	DETCRYS	Detector Scintillation Crystals	Bismuth Germanium Oxide	
2	ABC123	DO	44	1	DETCRYS	Detector Scintillation Crystals	Bismuth Germanium Oxide	
3	ABC123	DO	8842	1	TUBECOMP	Tube Composition	Polypropylene	

The RELREC table below uses --LNKID to relate the PR and AG domains to each other and to NV and --REFID to relate NV and DU.

In this example, the sponsor has maintained two sets of reference identifiers for the specific purpose of being able to relate records across multiple domains. Because the SDTMIG-MD advocates the use of --REFID to link a group of settings to the results obtained from the reading or interpretation of the test (see SDTMIG-MD v1.0, Section 4.2.1, Assumption 8), --LNKID has been used to establish the relationships between the procedure, the substance administered during the procedure, and the results obtained from the procedure. --LNKID is unique for each procedure for each subject, so datasets may be related to each other as a whole.

Rows 1-2: Show the relationship between the scan, represented in PR, and the radiolabel tracer used, represented in AG. There is only one tracer administration for each scan, and only one scan for each tracer administration, so the relationship is ONE to ONE.
Rows 3-4: Show the relationship between the scan, represented in PR, and the SUVR results obtained from the scan, represented in NV. Each scan yields two results, so the relationship is ONE to MANY.
Rows 5-6: Show the relationship between the radiolabel tracer used and the SUVR results for each scan. This relationship may seem indirect, but it is not: the choice of radiolabel has the potential to affect the results obtained. Because the relationship between PR and AG is ONE to ONE and the relationship between PR and NV is ONE to MANY, the relationship between AG and NV must be ONE to MANY.
Rows 7-8: Show the relationship between the SUVR results and the specific settings for the device used for each scan. There is more than one result from each scan, and more than one setting for each scan, so the relationship is MANY to MANY. This relationship is unusual and challenging to manage in a join/merge and only represents the concept of this relationship. To uniquely identify all the settings (DUTEST) related to an individual result (NVTEST), select an individual NVTEST and use its NVREFID to join/merge with the DUREFID to revise this to a ONE to MANY relationship to obtain all related DUTEST records.

relrec.xpt

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID
1	ABC123	PR		PRLNKID		ONE	6
2	ABC123	AG		AGLNKID		ONE	6
3	ABC123	PR		PRLNKID		ONE	7
4	ABC123	NV		NVLNKID		MANY	7
5	ABC123	AG		AGLNKID		ONE	8
6	ABC123	NV		NVLNKID		MANY	8
7	ABC123	NV		NVREFID		MANY	9
8	ABC123	DU		DUREFID		MANY	9

3.3 Scales of Cognitive Function

Many of the primary and secondary endpoints collected in trials of dementia conditions, such as AD and MCI, come in the form of clinical scales of cognition/function. These scales are typically represented in the Questionnaires (QS) Domain. The majority of the new documentation produced for this version of the Alzheimer's User Guide originates from Questionnaire supplements, maintained as standalone guides on the CDISC website at http://www.cdisc.org/content2909 .

The table below lists the scales that have been implemented in these supplemental user guides for AD as of the publication of this guide. Sponsors should refer to the link above if a scale of interest is not included below, as it may have been developed for another therapeutic area, and new scales are implemented in an ongoing basis by the QS Terminology and Standards Development subteams. See CDISC COP 017 CDISC SDTMIG Questionnaire Supplements (http://www.cdisc.org/bylaws-and-policies) for details on implementing or requesting development of standard questionnaires for SDTM-based submissions.

Full Name and Abbreviation	Status of Permission to develop controlled terminology	Status of controlled terminology	Status of supplement development
Clinical Dementia Rating Scale (CDR)†	Permission received	Terminology developed.	Supplement developed.
Geriatric Depression Scale –(GDS)	Permission received	Terminology developed.	Supplement developed.
Geriatric Depression Scale Short Form –(GDS-Short Version)	Permission received	Terminology developed.	Supplement developed.
Functional Activities Questionnaire-(FAQ)	Permission received	Terminology developed.	Supplement developed.
Functional Activities Questionnaire: NACC Version-(FAQ-NACC)	Permission received	Terminology developed.	Supplement developed.
Disability Assessment for Dementia-(DAD)	Permission received	Terminology developed.	Supplement developed.
Activities of Daily Living Inventory - MCI (ADCS-ADL MCI)	Permission received	Terminology developed.	Supplement developed.
Neuropsychiatric Inventory (NPI)	Permission received	Terminology developed.	Supplement developed.
Alzheimer's Disease Assessment Scale – Cognitive (ADAS-Cog)	Permission received	Terminology developed.	Supplement developed.
Mini-Mental State Examination (MMSE)	Permission denied	*	*
Clinical Global Impression Questionnaire (CGI)	Permission received	Terminology developed.	Supplement developed.
Modified Hachinski Ischemic Scale	Permission received	Terminology in development.	Supplement in development
Auditory Verbal Learning Test (AVLT)	Permission Received	Terminology Developed	Supplement in development

†Supplement is available on the CDISC website under the "Members Only" section.
**In general, CDISC does not develop supplements for questionnaires for which permission to develop controlled terminology has been denied. A supplement for MMSE has been developed, but the terminology it uses is generic.*

Appendices

Appendix A: Work Group

Core Development Team	
Name	**Institution/Organization**
Jon Neville	Critical Path Institute
Bob Stafford	Critical Path Institute
Bess LeRoy	Critical Path Institute
Steve Kopko	CDISC

Clinical Science and Technical Subject Matter Experts	
Name	**Institution/Organization**
Andreas Jeromin	NextGen Sciences
Jerry Novak	Johnson & Johnson
Pat Cole	Lilly
Mark Austin	Ixico
Karen Crawford	LONI/ADNI
Joyce Suhy	Synarc
Susan De Santi	GE
Diane Stephenson	Critical Path Institute/CAMD

Terminology and Questionnaires Supplement Development	
Name	**Institution/Organization**
Bernice Yost	CDISC
Erin Muhlbradt	NCI EVS
Roberta Rosenberg	Pfizer
Gary Cunningham	Teva
Jordan Li	NCI EVS
Donna Sattler	MaxisIT
Fred Wood	Accenture
Randall Austin	Glaxo-Smith Kline

Appendix B: Glossary and Abbreviations

AD	Alzheimer's Disease
CAMD	Coalition Against Major Diseases, a program of the Critical Path Institute
CDE	Common Data Element
CDISC	Clinical Data Interchange Standards Consortium
C-Path	Critical Path Institute
CRF	Case Report Form
CSF	Cerebrospinal Fluid
DICOM	Digital Imaging and Communications in Medicine
FDA	Food and Drug Administration
FDG	^{18}Fluorodeoxy Glucose- a radiolabeled tracer used in Positron Emission Tomography
MCI	Mild Cognitive Impairment
MRI	Magnetic Resonance Imaging
NCI EVS	National Cancer Institute Enterprise Vocabulary Services
PET	Positron Emission Tomography
PET/CT	Positron Emission Tomography-Computerized Tomography, a combination scanner using PET and CT
SDS	Submission Data Standards
SDTM	Study Data Tabulation Model
SDTMIG	Study Data Tabulation Model Implementation Guide for Human Clinical Trials
SDTMIG-MD	Study Data Tabulation Model Implementation Guide for Medical Devices
SME	Subject Matter Expert
SRC	Standards Review Council
SUVR	Standard Uptake Value Ratio

Appendix B1: Supplemental Qualifiers Name Codes

The following table contains additional standard QNAM and QLABEL values for use in the Supplemental Qualifiers (SUPP--) special-purpose datasets.

QNAM	QLABEL	Applicable Domains
MHOSTDTC	Onset of Symptoms Date	MH
SFTWR	Analysis Software	MO
SFTWRVER	Software Version	MO
REFREG	Reference Region	NV

Appendix C: Additional Assumptions for Domains

The domains used in this document operate under certain assumptions, which may be found with the domain specification tables in the implementation guides to which they belong (e.g., the SDTMIG for domains pertaining to human clinical trials, the SDTMIG-MD for domains pertaining to the use of medical devices, the SDTM APIG for domains pertaining to persons who are relevant to the study without being study subjects). All of the assumptions from those implementation guides apply for this user guide. The assumptions given below apply to TAUG-Alzheimer's ***in addition*** to the standard assumptions.

Appendix C1: General Assumptions for AD/MCI

1. Controlled terminology is still under development for the AD data standards, thus some values in the examples are not CDISC controlled terms. Sponsors should always check terminology shown in the examples against current controlled terminology before adopting it.

Appendix C2: Additional Assumptions for Procedure Agents Domain Model

1. AGLNKID ties the radiopharmaceutical to the imaging procedure in the PR domain.
2. It is important to know the time elapsed between tracer injection and the start of the scan. Therefore, it is important to record AGSTDTC to the level of minutes, or even seconds. Time elapsed can then be determined by the difference in AGSTDTC and PRSTDTC (scan start time).

Appendix C3: Additional Assumptions for Procedures Domain Model

1. Sponsors should use the PR domain to represented procedures such as lumbar puncture and imaging, including MRI or PET scans, if such data was collected. Subsequent handling of the CSF specimen obtained via lumbar puncture (aliquoting, freezing, thawing) should be represented in the Biospecimen domains (BE and BS), if collected. Any results related to CSF samples or imaging should be represented in the appropriate findings domains (LB, NV, or MO), as shown in the domain examples.
2. PRLNKID is used to connect the imaging procedure with scan results. In the case of PET imaging, it is also used to link to the radiolabeled tracer information.
3. PRREFID is equal to BEREFID when BECAT=COLLECTION.

Appendix C4: Additional Assumptions for Medical History Domain Model

1. "Alzheimer's disease", "mild cognitive impairment", or other terms relating to Alzheimer's-type dementias for which the study is being conducted should be represented in MHTERM when MHCAT=Primary Diagnosis. Coding to the Preferred Term with MedDRA for Alzheimer's disease or mild cognitive impairment to populate MHDECOD is optional but encouraged. See Section 2.2 for how to map the primary diagnosis.
2. MHSTDTC should be populated with the date of diagnosis when MHCAT=PRIMARY DIAGNOSIS.
3. If general medical history information is also captured either as verbatim text or pre-specified text, it can be coded with MedDRA and the Preferred Term reported in MHDECOD based on the sponsor's coding criteria for medical history.
4. Terminology:
    a. The terminology for MHCAT includes "PRIMARY DIAGNOSIS and "GENERAL."
    b. Note that MHCAT is not subject to controlled terminology, and this is just a recommendation.
    c. Where collected, onset date of symptoms for either Alzheimer's disease or MCI should be mapped to Supplemental Qualifiers (SUPPMH).
    d. Existing CDISC Controlled Terminology should be used for the MH domain. There is no additional CDISC Controlled Terminology needed for the MH domain for AD.

Appendix C5: Additional Assumptions for Laboratory Test Results Domain Model

1. If lab results were obtained from an aliquot derived from the original sample, the unique identifier for the aliquot (value in BSREFID/BEREFID), if important, should be represented in LBREFID as well.
2. If a device (such as a 96-well plate) is directly used to obtain a lab result, SPDEVID should be populated to match the SPDEVID of the corresponding device in the DI domain.

Appendix C6: Additional Assumptions for Morphology Domain Model

1. SPDEVID identifies the MRI scanner used to acquire these measures. It should be matched with the same SPDEVID in the device domains.
2. MOREFID links the morphology result to the particular scan and the properties/settings of the scan in the Device In-Use domain (DUREFID).

3. MOLNKID links the morphology result to the image acquisition procedure in the PR domain (PRLNKID).
4. When the method used to determine the results shown is SOFTWARE ANALYSIS (as in the example in Section 3.2.1.2), the software used to analyze the images should be represented in SUPPMO. The software name and version should be linked to the subject identifier.

Appendix C7: Additional Assumptions for Nervous System Findings Domain Model

1. Sponsors should represent as many target brain regions for which they calculated SUVR values, using one SUVR test per row and showing each associated brain region in NVLOC. Associated qualifiers of NVLOC (NVLAT, NVDIR) may be used where appropriate. The examples shown in Section 3.2.2.2 are not meant to be all inclusive. Each reference region used for SUVR determination should be listed in the SUPPNV domain, linking back to the SUVR record to which it applies.
2. NVREFID relates SUVR values obtained from an imaging scan to the scan properties/settings in the DU domain (DUREFID).
3. NVLNKID links the result in NV to the radiolabeled tracer used in the AG domain to (AGLNKID). It is important to maintain this link as it defines the tracer used, and therefore the target biomarker being imaged. It also links the result to the imaging procedure in the PR domain (PRLNKID).
4. SPDEVID shows the scanner ID number used to generate the images from which these results were obtained.

Appendix C8: Additional Assumptions for Pharmacogenomic/Genetics Domains

1. BE domain:
    a. Information about Biospecimen Events (such as collection, aliquoting, freezing) is represented in detail in the BE domain. Since each sample needs a separate identifier to link it to further actions or characteristics of the sample, each aliquoting event is assigned a unique BEREFID value that can be traced to the BEREFID assigned for the collected "parent" sample. BEREFID is used to connect the BE and BS (BSREFID) domains, as well as to any results in the LB domain (LBREFID) obtained from the sample.
2. BS domain:
    a. SPDEVID identifies the device associated with any biospecimen property. It links the BS domain with the device domains.
    b. BSREFID links the listed findings of the biospecimen to any biospecimen events by linking to BEREFID.
3. PF domain:
    a. If genotyping information is collected it should be represented in the Pharmacogenomics Findings domain.

Appendix C9: Additional Assumptions for Medical Device Domains

1. DI domain: The information for tube type in this domain refers to a specific lot of tubes, since the properties within a lot do not change. Properties of tubes were identified as being important to interpreting results of biomarker tests. In cases where the tube properties or the properties of any other "devices" are not deemed important to interpreting results, they do not need to be represented.

Appendix C10: Additional Assumptions for Associated Persons Domains

1. AD special-interest terminology for the following Associated Persons domains is displayed in the examples. Note that controlled terminology is currently out-of-scope for MHCAT and MHTERM. All other controlled terminology should be used as appropriate.

a. Medical History: MHCAT and MHTERM use AD special-interest terminology
    i. MHCAT=PRIMARY DIAGNOSIS
    ii. MHTERM= Alzheimer's disease or Mild Cognitive Impairment

Appendix D: Representations and Warranties, Limitations of Liability, and Disclaimers

CDISC Patent Disclaimers
It is possible that implementation of and compliance with this standard may require use of subject matter covered by patent rights. By publication of this standard, no position is taken with respect to the existence or validity of any claim or of any patent rights in connection therewith. CDISC, including the CDISC Board of Directors, shall not be responsible for identifying patent claims for which a license may be required in order to implement this standard or for conducting inquiries into the legal validity or scope of those patents or patent claims that are brought to its attention.

Representations and Warranties
"CDISC grants open public use of this User Guide (or Final Standards) under CDISC's copyright."

Each Participant in the development of this standard shall be deemed to represent, warrant, and covenant, at the time of a Contribution by such Participant (or by its Representative), that to the best of its knowledge and ability: (a) it holds or has the right to grant all relevant licenses to any of its Contributions in all jurisdictions or territories in which it holds relevant intellectual property rights; (b) there are no limits to the Participant's ability to make the grants, acknowledgments, and agreements herein; and (c) the Contribution does not subject any Contribution, Draft Standard, Final Standard, or implementations thereof, in whole or in part, to licensing obligations with additional restrictions or requirements inconsistent with those set forth in this Policy, or that would require any such Contribution, Final Standard, or implementation, in whole or in part, to be either: (i) disclosed or distributed in source code form; (ii) licensed for the purpose of making derivative works (other than as set forth in Section 4.2 of the CDISC Intellectual Property Policy ("the Policy")); or (iii) distributed at no charge, except as set forth in Sections 3, 5.1, and 4.2 of the Policy. If a Participant has knowledge that a Contribution made by any Participant or any other party may subject any Contribution, Draft Standard, Final Standard, or implementation, in whole or in part, to one or more of the licensing obligations listed in Section 9.3, such Participant shall give prompt notice of the same to the CDISC President who shall promptly notify all Participants.

No Other Warranties/Disclaimers. ALL PARTICIPANTS ACKNOWLEDGE THAT, EXCEPT AS PROVIDED UNDER SECTION 9.3 OF THE CDISC INTELLECTUAL PROPERTY POLICY, ALL DRAFT STANDARDS AND FINAL STANDARDS, AND ALL CONTRIBUTIONS TO FINAL STANDARDS AND DRAFT STANDARDS, ARE PROVIDED "AS IS" WITH NO WARRANTIES WHATSOEVER, WHETHER EXPRESS, IMPLIED, STATUTORY, OR OTHERWISE, AND THE PARTICIPANTS, REPRESENTATIVES, THE CDISC PRESIDENT, THE CDISC BOARD OF DIRECTORS, AND CDISC EXPRESSLY DISCLAIM ANY WARRANTY OF MERCHANTABILITY, NONINFRINGEMENT, FITNESS FOR ANY PARTICULAR OR INTENDED PURPOSE, OR ANY OTHER WARRANTY OTHERWISE ARISING OUT OF ANY PROPOSAL, FINAL STANDARDS OR DRAFT STANDARDS, OR CONTRIBUTION.

Limitation of Liability
IN NO EVENT WILL CDISC OR ANY OF ITS CONSTITUENT PARTS (INCLUDING, BUT NOT LIMITED TO, THE CDISC BOARD OF DIRECTORS, THE CDISC PRESIDENT, CDISC STAFF, AND CDISC MEMBERS) BE LIABLE TO ANY OTHER PERSON OR ENTITY FOR ANY LOSS OF PROFITS, LOSS OF USE, DIRECT, INDIRECT, INCIDENTAL, CONSEQUENTIAL, OR SPECIAL DAMAGES, WHETHER UNDER CONTRACT, TORT, WARRANTY, OR OTHERWISE, ARISING IN ANY WAY OUT OF THIS POLICY OR ANY RELATED AGREEMENT, WHETHER OR NOT SUCH PARTY HAD ADVANCE NOTICE OF THE POSSIBILITY OF SUCH DAMAGES.

Note: The CDISC Intellectual Property Policy can be found at http://www.cdisc.org/about/bylaws_pdfs/CDISCIPPolicy-FINAL.pdf.

6 Domain Models Based on the General Observation Classes

6.1 Interventions

AG – Procedure Agents

Some tests involve administration of substances, and it has been unclear what domain these should be stored in. The Concomitant Medications domain seemed particularly inappropriate when the substance was one that would never been given as a medication. Even substances that are medications are not being used as such when they are given as part of a testing procedure. The Exposure domain also seemed inappropriate, since although the testing procedure might be part of the study plan, these data would not be used or analyzed in the same way as data about study treatments. The Procedure Agents domain was created to fill this gap. The Procedure Agents domain has advantages over the draft Procedures domain for this purpose. It allows recording of multiple substance administrations for a single testing procedure. It also separates data about substance administrations from data about procedures which do not involve substance administration.

AG – Description/Overview for Procedure Agents Domain Model

The Procedure Agents domain is a draft domain at the time of this publication. No CDISC controlled terminology definition exists for the domain yet.

Both the provisional Procedures domain and this draft Procedure Agents domain allow collection of doses administered during a procedure, and discussions are ongoing to provide guidance on deciding what data should be stored in which domain. The draft Procedure Agents domain can be used to provide data on several substance administrations within the same procedure, as shown in Example 2 below.

AG – Specification for Procedure Agents Domain Model

ag.xpt, Procedure Agents — Interventions, Version 3.x.x. One record per recorded intervention occurrence per subject, Tabulation.

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Role	CDISC Notes	Core
STUDYID	Study Identifier	Char		Identifier	Unique identifier for a study.	Req
DOMAIN	Domain Abbreviation	Char	AG	Identifier	Two-character abbreviation for the domain.	Req
USUBJID	Unique Subject Identifier	Char		Identifier	Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product.	Req
AGSEQ	Sequence Number	Num		Identifier	Sequence Number given to ensure uniqueness of subject records within a domain. May be any valid number.	Req
AGGRPID	Group ID	Char		Identifier	Used to tie together a block of related records in a single domain for a subject.	Perm

AGSPID	Sponsor-Defined Identifier	Char		Identifier	Sponsor-defined reference number. Perhaps pre-printed on the CRF as an explicit line identifier or defined in the sponsor's operational database. Example: Line number from the procedure or test page.	Perm
AGTRT	Reported Agent Name	Char		Topic	Verbatim medication name that is either pre-printed or collected on a CRF.	Req
AGMODIFY	Modified Reported Name	Char		Synonym Qualifier	If AGTRT is modified to facilitate coding, then AGMODIFY will contain the modified text.	Perm
AGDECOD	Standardized Agent Name	Char	*	Synonym Qualifier	Standardized or dictionary-derived text description of AGTRT or AGMODIFY. Equivalent to the generic medication name in WHO Drug. The sponsor is expected to provide the dictionary name and version used to map the terms utilizing the define.xml external codelist attributes. If an intervention term does not have a decode value in the dictionary then AGDECOD will be left blank.	Perm
AGCAT	Category for Agent	Char	*	Grouping Qualifier	Used to define a category of agent. Examples: CHALLENGE AGENT, or PET TRACER.	Perm
AGSCAT	Subcategory for Agent	Char	*	Grouping Qualifier	Further categorization of agent.	Perm
AGPRESP	AG Pre-Specified	Char	(NY)	Record Qualifier	Used to indicate whether (Y/null) information about the use of a specific agent was solicited on the CRF.	Perm
AGOCCUR	AG Occurrence	Char	(NY)	Record Qualifier	When the use of specific agent is solicited, AGOCCUR is used to indicate whether or not (Y/N) use of the agent occurred. Values are null for agents not specifically solicited.	Perm
AGSTAT	Completion Status	Char	(ND)	Record Qualifier	Used to indicate that a question about a pre-specified agent was not answered. Should be null or have a value of NOT DONE.	Perm
AGREASND	Reason Test Not Performed	Char		Record Qualifier	Describes the reason procedure agent was not collected. Used in conjunction with AGSTAT when value is NOT DONE.	Perm
AGCLAS	Agent Class	Char	*	Variable Qualifier	Drug class. May be obtained from coding. When coding to a single class, populate with class value. If using a dictionary and coding to multiple classes, then follow assumption 4.1.2.8.3 or omit AGCLAS.	Perm
AGCLASCD	Agent Class Code	Char	*	Variable Qualifier	Class code corresponding to AGCLAS. Drug class. May be obtained from coding. When coding to a single class, populate with class code. If using a dictionary and coding to multiple classes, then follow assumption 4.1.2.8.3 or omit AGCLASCD.	Perm
AGDOSE	Dose per Administration	Num		Record Qualifier	Amount of AGTRT taken.	Perm
AGDOSTXT	Dose Description	Char		Record Qualifier	Dosing amounts or a range of dosing information collected in text form. Units may be stored in AGDOSU. Example: 200-400, 15-20.	Perm
AGDOSU	Dose Units	Char	(UNIT)	Variable Qualifier	Units for AGDOSE and AGDOSTXT. Examples: ng, mg, or mg/kg.	Perm
AGDOSFRM	Dose Form	Char	(FRM)	Variable Qualifier	Dose form for AGTRT. Examples: TABLET, AREOSOL.	Perm
AGDOSFRQ	Doing Frequency per Interval	Char	(FREQ)	Variable Qualifier	Usually expressed as the number of repeated administrations of AGDOSE within a specific time period. Example: ONCE	Perm
AGROUTE	Route of Administration	Char	(ROUTE)	Variable Qualifier	Route of administration for AGTRT. Examples: ORAL.	Perm

VISITNUM	Visit Number	Num		Timing	1. Clinical encounter number. 2. Numeric version of VISIT, used for sorting.	Exp
VISIT	Visit Name	Char		Timing	1. Protocol-defined description of clinical encounter. 2. May be used in addition to VISITNUM and/or VISITDY.	Perm
VISITDY	Planned Study Day of Visit	Num		Timing	Planned study day of the visit based upon RFSTDTC in Demographics.	Perm
AGSTDTC	Start Date/Time of Agent	Char	ISO 8601	Timing		Perm
AGENDTC	End Date/Time of Agent	Char	ISO 8601	Timing		Perm
AGSTDY	Study Day of Start of Agent	Num		Timing	Study day of start of agent relative to the sponsor-defined RFSTDTC.	Perm
AGENDY	Study Day of End of Agent	Num		Timing	Study day of end of agent relative to the sponsor-defined RFSTDTC.	Perm
AGDUR	Duration of Agent	Char	ISO 8601	Timing	Collected duration for an agent episode. Used only if collected on the CRF and not derived from start and end date/times.	Perm
AGSTRF	Start Relative to Reference Period	Char	(STENRF)	Timing	Describes the start of the agent relative to sponsor-defined reference period. The sponsor-defined reference period is a continuous period of time defined by a discrete starting point and a discrete ending point (represented by RFSTDTC and RFENDTC in Demographics). If information such as "PRIOR", "ONGOING", or "CONTINUING" was collected, this information may be translated into AGSTRF.	Perm
AGENRF	End Relative to Reference Period	Char	(STENRF)	Timing	Describes the end of the agent relative to the sponsor-defined reference period. The sponsor-defined reference period is a continuous period of time defined by a discrete starting point and a discrete ending point (represented by RFSTDTC and RFENDTC in Demographics). If information such as "PRIOR", "ONGOING", or "CONTINUING" was collected, this information may be translated into AGENRF.	Perm
AGSTRTPT	Start Relative to Reference Time Point	Char	BEFORE, COINCIDENT, AFTER, U	Timing	Identifies the start of the agent as being before or after the reference time point defined by variable AGSTTPT.	Perm
AGSTTPT	Start Reference Time Point	Char		Timing	Description or date/time in ISO 8601 character format of the reference point referred to by AGSTRTPT. Examples: "2003-12-15" or "VISIT 1".	Perm
AGENRTPT	End Relative to Reference Time Point	Char	BEFORE, COINCIDENT, AFTER, ONGOING, U	Timing	Identifies the end of the agent as being before or after the reference time point defined by variable AGENTPT.	Perm
AGENTPT	End Reference Time Point	Char		Timing	Description or date/time in ISO 8601 character format of the reference point referred to by AGENRTPT. Examples: "2003-12-25" or "VISIT 2".	Perm

** Indicates variable may be subject to controlled terminology, (Parenthesis indicates CDISC/NCI codelist code value)*

AG – Assumptions for Procedure Agents Domain Model

1. AG Definition and Structure
    a. CRF data that captures the agents administered to the subject as part of a procedure or assessment as opposed to drugs, medications and therapies administered with therapeutic intent. An example is a short-acting bronchodilator administered as part of a reversibility assessment. Other examples of substance administrations that could be submitted in this domain include contrast agents and radio labeled substances used in imaging studies. Discussions are ongoing on the handling of radiation (e.g., x-rays or visible light) in SDTM interventions domains.

   b. The structure of the AG domain is one record per agent intervention episode, or pre-specified agent assessment per subject. It is the sponsor's responsibility to define an intervention episode. This definition may vary based on the sponsor's requirements for review and analysis.
2. Procedure Agent Description and Coding
   a. AGTRT captures the name of the agent and it is the topic variable. It is a required variable and must have a value. AGTRT should include only the agent name, and should not include dosage, formulation, or other qualifying information. For example, ALBUTEROL 2 PUFF is not a valid value for AGTRT. This example should be expressed as AGTRT = ALBUTEROL, AGDOSE = 2, AGDOSU = PUFF, and AGDOSFRM = AEROSOL
   b. AGMODIFY should be included if the sponsor‘s procedure permits modification of a verbatim term for coding.
   c. AGDECOD is the standardized agent term derived by the sponsor from the coding dictionary. It is possible that the reported term (AGTRT) or the modified term (AGMODIFY) can be coded using a standard dictionary. In this instance the sponsor is expected to provide the dictionary name and version used to map the terms utilizing the define.xml external codelist attributes.
3. Pre-specified Terms; Presence or Absence of Procedure Agents
   a. AGPRESP is used to indicate whether an agent was pre-specified.
   b. AGOCCUR is used to indicate whether a pre-specified agent was used. A value of Y indicates that the agent was used and N indicates that it was not.
   c. If an agent was not pre-specified the value of AGOCCUR should be null. AGPRESP and AGOCCUR are permissible fields and may be omitted from the dataset if all agents were collected as free text. Values of AGOCCUR may also be null for pre-specified agents if no Y/N response was collected; in this case, AGSTAT = NOT DONE, and AGREASND could be used to describe the reason the answer was missing.
4. Additional Permissible Interventions Qualifiers
   a. The variables --INDC, --DOSTOT, and --DOSRGM from the Interventions general observation class would not generally be used in the AG domain because AG should only contain agents used as part of a procedure or an assessment.
   b. Other additional Qualifiers from the SDTM Interventions Class may be added to this domain.

AG – Examples for Procedure Agents Domain Model

Example 1

This example shows the administration of a procedure agent administered as part of a reversibility assessment with the associated spirometer results, as well as the spirometry measurements (RE domain) obtained before and after agent administration. Depending on the study design, the route of bronchodilator administration (via meter dose inhaler (MDI) or nebulizer) and dose per actuation (puff) or nebule may also be collected.

Reversibility Assessment

Date of assessment: DD-MMM-YYYY
Was the subject administered a short-acting bronchodilator in the previous 4 hours? Yes No
Pre-Bronchdilator Spirometry (5 Minutes before Albuterol Dosing)
    Time of Assessment: HH:MM
    Forced Expiratory Volume in 1 Second (FEV1) Result: _____L
Albuterol Administration
    Was the subject administered Albuterol? Yes No
    Time of Assessment: HH:MM
    Number of Puffs administered: ____

Post-Bronchdilator Spirometry (20 Minutes after Albuterol Dosing)
Time of Assessment: HH:MM
Forced Expiratory Volume in 1 Second (FEV1) Result: _____L
Percentage Reversibility: _____ %

Row 1: Shows the administration data of an agent (Albuterol) which was pre-specified on the CRF as part of the reversibility procedure.

ag.xpt

Row	STUDYID	DOMAIN	USUBJID	AGSEQ	AGTRT	AGPRESP	AGOCCUR	AGDOSE	AGDOSU	AGDOSFRM	AGDOSFRQ	AGROUTE	VISIT	AGSTDTC
1	XYZ	AG	XYZ-001-001	1	ALBUTEROL	Y	Y	2	PUFF	AEROSOL	ONCE	ORAL	VISIT 2	2013-06-18T10:05

Row 1: Shows the record where the question as to whether a short-acting bronchodilator was administered in the 4 hours prior to the reversibility assessment. A short-acting bronchodilator administered prior to the reversibility test, is used with therapeutic intent so is tabulated in the CM domain. Note that AGTRT has been populated with a description of a kind of medication rather than a single medication.

cm.xpt

Row	STUDYID	DOMAIN	USUBJID	CMSEQ	CMTRT	CMPRESP	CMOCCUR	CMEVLINT
1	XYZ	CM	XYZ-001-001	1	SHORT-ACTING BRONCHODILATOR	Y	N	-PT4H

Row 1: Shows the data in original and standardized units of measure in REORRES, RESTRESC and RESTRESN for FEV1 of a pre-bronchodilator-administration spirometry test performed as part of a reversibility assessment with the associated timing reference variables RETPT, RETPTNUM, REELTM, RETPTREF, and RERFTDTC. This test was performed 5 minutes before the bronchodilator challenge.

Row 2: Shows the data in original and standardized units of measure in REORRES, RESTRESC and RESTRESN for FEV1 of a post-bronchodilator administration spirometry test performed as part of a reversibility assessment with the associated timing reference variables RETPT, RETPTNUM, REELTM, RETPTREF, and RERFTDTC. This test was performed 20 minutes after the bronchodilator challenge.

Row 3: Shows the data in original and standardized units of measure in REORRES, RESTRESC and RESTRESN for the percentage reversibility where this is collected.

re.xpt

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	RESEQ	REGRPID	RETESTCD	RETEST	REORRES	REORRESU	RESTRESC	RESTRESN
1	XYZ	RE	XYZ-001-001	ABC001	1	1	FEV1	Forced Expiratory Volume in 1 Second	2.43	L	2.43	2.43
2	XYZ	RE	XYZ-001-001	ABC001	2	1	FEV1	Forced Expiratory Volume in 1 Second	2.77	L	2.77	2.77
3	XYZ	RE	XYZ-001-001	ABC001	3	1	PCTREV	Percentage Reversibility	13.99	%	13.99	13.99

Row	RESTRESU	VISIT	REDTC	RETPT	RETPTNUM	REELTM	RETPTREF	RERFTDTC
1 (cont)	L	VISIT 2	2013-06-18T10:00	PRE-BRONCHODILATOR ADMINISTRATION	1	-PT5M	BRONCHODILATOR ADMINISTRATION	2013-06-18T10:05
2 (cont)	L	VISIT 2	2013-06-18T10:25	POST-BRONCHODILATOR ADMINISTRATION	2	PT20M	BRONCHODILATOR ADMINISTRATION	2013-06-18T10:05
3 (cont)	%	VISIT 2	2013-06-18T10:25				BRONCHODILATOR ADMINISTRATION	2013-06-18T10:05

Row 1: Shows the device type that was used for the pulmonary function tests as part of the reversibility procedure.

di.xpt

Row	STUDYID	DOMAIN	SPDEVID	DISEQ	DIPARMCD	DIPARM	DIVAL
1	XYZ	DI	ABC001	1	TYPE	Device Type	SPIROMETER

Rows 1-3: Shows the relationship of the test agent to the spirometry measurements obtained before and after its administration and to the prior occurrence of short acting bronchodilator administration.

relrec.xpt

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID
1	XYZ	AG	XYZ-001-001	AGSEQ	1		1
2	XYZ	RE	XYZ-001-001	REGRPID	1		1
3	XYZ	CM	XYZ-001-001	CMSEQ	1		1

Example 2

This example captures data about the allergen used by the subject as part of a bronchial allergen challenge (BAC) test. Initially, the subject had a skin prick allergen test to help identify the allergen to be used for the BAC test. The allergens tested were cat dander, house dust mite, and grass. For this subject, grass provided the largest skin test reaction and was the allergen chosen to be used in the BAC test. A predetermined set of ascending doses of the chosen allergen are used in the screening BAC test. The results of the screening BAC are used to choose the allergen dose that will be used in subsequent BAC tests (not shown).

Allergen Used?
☐ Cat Dander
☐ House Dust Mites
☐ Grass

Inhalation End Time	Allergen Concentration SQ-u/mL		Time of FEV1	FEV1 (L)
___:___	Saline=0	0	___:___	__.__
___:___	Dose1	250	___:___	__.__
___:___	Dose2	1000	___:___	__.__
___:___	Dose3	2000	___:___	__.__

Rows 1-3: Correspond to the first part of the CRF. The skin response results corresponding to these allergen administrations were used to choose grass as the allergen for the BAC.
Rows 4: The first dose given in the BAC was saline.
Rows 5-6: Three successively higher doses of grass allergen were given.

ag.xpt

Row	STUDYID	DOMAIN	USUBJID	AGSEQ	AGTRT	AGPRESP	AGOCCUR	AGDOSE	AGDOSU	AGROUTE	VISIT	AGENDTC
1	XYZ	AG	XYZ-001-001	1	CAT DANDER	Y	N			INTRAEPIDERMAL	SCREENING	2010-10-31
2	XYZ	AG	XYZ-001-001	2	HOUSE MITE DUST	Y	N			INTRAEPIDERMAL	SCREENING	2010-10-31
2	XYZ	AG	XYZ-001-001	3	GRASS	Y	Y			INTRAEPIDERMAL	SCREENING	2010-10-31
3	XYZ	AG	XYZ-001-001	4	SALINE	Y	Y	0	SQ-u/mL	RESPIRATORY (INHALATION)	SCREENING	2010-11-07T10:56:00
4	XYZ	AG	XYZ-001-001	5	GRASS	Y	Y	250	SQ-u/mL	RESPIRATORY (INHALATION)	SCREENING	2010-11-07T11:19:00
5	XYZ	AG	XYZ-001-001	6	GRASS	Y	Y	1000	SQ-u/mL	RESPIRATORY (INHALATION)	SCREENING	2010-11-07T11:43:00
6	XYZ	AG	XYZ-001-001	7	GRASS	Y	Y	2000	SQ-u/mL	RESPIRATORY (INHALATION)	SCREENING	2010-11-07T12:06:00

3 Biospecimen Domains

3.1 Biospecimen Events (BE)

3.1.1 Description/Overview for Biospecimen Events Domain Model

The BE domain has no controlled terminology definiton yet.

--PARENT and --SPCLVL are variables for recording the hierarchy of specimen relationships. There is some question as to whether they belong in BE, BS, or as part of a special purpose relationships dataset (RELSPEC). For now, all three options are given.

3.1.2 Specification for Biospecimen Events Domain Model

be.xpt, Biospecimen Events — Events. One record per biospecimen event per specimen collected per subject, Tabulations

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Role	CDISC Notes	Core
STUDYID	Study Identifier	Char		Identifier	Unique identifier for a study.	Req
DOMAIN	Domain Abbreviation	Char	BE	Identifier	Two-character abbreviation for the domain.	Req
USUBJID	Unique Subject Identifier	Char		Identifier	Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product.	Req
SPDEVID	Sponsor Device Identifier	Char		Identifier	Sponsor-defined identifier for a device.	Perm
BESEQ	Sequence Number	Num		Identifier	Sequence Number given to ensure uniqueness of subject records within a domain. May be any valid number.	Req
BEGRPID	Group ID	Char		Identifier	Used to tie together a block of related records in a single domain.	Perm
BEREFID	Reference ID	Char		Identifier	Optional internal or external identifier such as lab specimen ID.	Perm
BESPID	Sponsor-Defined Identifier	Char		Identifier	Optional sponsor-defined reference number. Perhaps pre-printed on the CRF as an explicit line identifier or defined in the sponsor's operational database. Example: Line number on a CRF page.	Perm
BETERM	Reported Term for the Biospecimen Event	Char		Topic	Topic variable for an event observation, which is the verbatim or pre-specified name of the event.	Req
BEMODIFY	Modified Reported Term	Char		Synonym Qualifier	If BETERM is modified to facilitate coding, then BEMODIFY will contain the modified text.	Perm
BEDECOD	Dictionary-Derived Term	Char	*	Synonym Qualifier	Dictionary-derived text description of BETERM or BEMODIFY. CDISC vocabulary will be defined.	Exp
BEPARTY	Accountable Party	Char	*	Record Qualifier	Person or organization accountable for the specimen at the conclusion of the action specified in BETERM.	Perm
BEPRTYID	Identification of specific accountable party	Char		Record Qualifier	An identifier for the responsible group/role (e.g. site, subject).	Perm

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Role	CDISC Notes	Core
BECAT	Category for Biospecimen Event	Char	*	Grouping Qualifier	Used to define a category of related records. Example: COLLECTION, PREP, TRANSPORT	Perm
BESCAT	Subcategory for Biospecimen Event	Char	*	Grouping Qualifier	A further categorization of biospecimen event. Example: For specimen collection, this may specify SOFT TISSUE.	Perm
BEBODSYS	Body System or Organ Class	Char	*	Record Qualifier	Body system or organ class used by the sponsor from the coding dictionary (e.g., MedDRA). When using a multi-axial dictionary such as MedDRA, this should contain the SOC used for the sponsor's analyses and summary tables which may not necessarily be the primary SOC.	Perm
BELOC	Body Location	Char	(LOC)	Record Qualifier	Describes body location relevant for the event (e.g. BRAIN, LUNG).	Perm
BEPARENT	Specimen Parent	Char		Record Qualifier	Identifies the parent of a specimen to support tracking its genealogy.	Perm
BESPCLVL	Specimen Level	Num		Record Qualifier	Identifies the generation number of the sample where the collected sample is considered the first generation.	Perm
VISITNUM	Visit Number	Num		Timing	1. Clinical encounter number. 2. Numeric version of VISIT, used for sorting.	Exp
VISIT	Visit Name	Char		Timing	1. Protocol-defined description of clinical encounters. 2. May be used in addition to VISITNUM and/or VISITDY	Perm
VISITDY	Planned Study Day of Visit	Num		Timing	Planned study day of the visit based upon RFSTDTC in Demographics.	Perm
BEDTC	Date/Time of Specimen Collection	Char	ISO 8601	Timing	Date and time of specimen collection.	Exp
BESTDTC	Start Date/Time of Biospecimen Event	Char	ISO 8601	Timing	Start Date/Time of a biospecimen event.	Exp
BEENDTC	End Date/Time of Biospecimen Event	Char	ISO 8601	Timing	End Date/Time of a biospecimen event	Exp
BESTDY	Study Day of Start of Biospecimen Event	Num		Timing	Study day of start of biospecimen event relative to the sponsor-defined RFSTDTC.	Perm
BEENDY	Study Day of End of Biospecimen Event	Num		Timing	Study day of end of event relative to the sponsor-defined RFSTDTC.	Perm
BEDUR	Duration of Bio specimen Event	Char	ISO 8601	Timing	Collected duration and unit of a biospecimen event. Used only if collected on the CRF and not derived from start and end date/times. Example: P1DT2H (for 1 day, 2 hours).	Perm

** Indicates variable may be subject to controlled terminology, (Parenthesis indicates CDISC/NCI codelist code value)*

3.1.3 Assumptions for Biospecimen Events domain model

1. The BE domain contains data about actions taken that affect or may affect a specimen, such as specimen collection, freezing and thawing, aliquoting, and transportation.
2. Event timing consists primarily of start/end date/times and/or durations.

3.2 Biospecimen (BS)

3.2.1 Description/Overview for Biospecimen Domain Model

The BS domain has no controlled terminology definition yet.

--PARENT and --SPCLVL are variables for recording the hierarchy of specimen relationships. There is some question as to whether they belong in BE, BS, or as part of a special purpose relationships dataset (RELSPEC). For now, all three options are given.

3.2.2 Specification for Biospecimen Domain Model

bs.xpt, Biospecimen— Findings. One record per bio-specimen finding per specimen collected per subject, Tabulation

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Role	CDISC Notes	Core
STUDYID	Study Identifier	Char		Identifier	Unique identifier for a study.	Req
DOMAIN	Domain Abbreviation	Char	BS	Identifier	Two-character abbreviation for the domain.	Req
USUBJID	Unique Subject Identifier	Char		Identifier	Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product.	Req
BSSEQ	Sequence Number	Num		Identifier	Sequence Number given to ensure uniqueness of subject records within a domain. May be any valid number.	Req
BSGRPID	Group ID	Char		Identifier	Used to tie together a block of related records in a single domain for a subject.	Perm
BSREFID	Reference ID	Char		Identifier	Optional internal or external identifier such as lab specimen ID.	Perm
SPDEVID	Sponsor Device Identifier	Char		Identifier	Sponsor-defined identifier for a device.	Perm
BSSPID	Sponsor-Defined Identifier	Char		Identifier	Sponsor-defined reference number. Perhaps pre-printed on the CRF as an explicit line identifier or defined in the sponsor's operational database. This may also be used to designate a sample derived from the collected specimen: Example: RNA or DNA extraction, tissue resectioning	Perm
BSTESTCD	Biospecimen Test or Examination Short Name	Char	*	Topic	Short name of the measurement, test, observation or examination described in BSTEST. It can be used as a column name when converting a dataset from a vertical to a horizontal format. The value in BSTESTCD cannot be longer than 8 characters, nor can it start with a number (e.g.'1TEST'). BSTESTCD cannot contain characters other than letters, numbers, or underscores. Examples**:** TISSTY, FLSFRZDT, ADDTIVOL	Req
BSTEST	Biospecimen Test or Examination Name	Char	*	Synonym Qualifier	Verbatim name of the test, observation or examination used to obtain the measurement or finding. Note any test normally performed by a clinical laboratory is considered a lab test. The value in BSTEST cannot be longer than 40 characters. Examples**:** Flash Frozen Date Time, Flash Frozen Duration, Additive Volume	Req
BSCAT	Category for Biospecimen Test	Char	*	Grouping Qualifier	Used to define a category of related records across subjects. Example: SPECIMEN HANDLING, SPECIMEN MEASUREMENT	Exp

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Role	CDISC Notes	Core
BSSCAT	Subcategory for Biospecimen Test	Char	*	Grouping Qualifier	A further categorization of a test category.	Perm
BSORRES	Result or Finding in Original Units	Char		Result Qualifier	Result of the measurement or finding as originally received or collected.	Exp
BSORRESU	Original Units	Char	*	Variable Qualifier	Original units in which the data were collected. The unit for BSORRES. Examples: mg. or ml.	Exp
BSSTRESC	Character Result/Finding in Standard Format	Char		Result Qualifier	Contains the result value for all findings, copied or derived from BSORRES in a standard format or standard units. BSSTRESC should store all results or findings in character format; if results are numeric, they should also be stored in numeric format in BSSTRESN. For example, if a test has results 'NONE', 'NEG', and 'NEGATIVE' in BSORRES and these results effectively have the same meaning, they may be represented in standard format in BSSTRESC as 'NEGATIVE'.	Exp
BSSTRESN	Numeric Result/Finding in Standard Units	Num		Result Qualifier	Used for continuous or numeric results or findings in standard format; copied in numeric format from BSSTRESC. BSSTRESN should store all numeric test results or findings.	Exp
BSSTRESU	Standard Units	Char	*	Variable Qualifier	Standardized unit used for BSSTRESC or BSSTRESN.	Exp
BSSTAT	Completion Status	Char	(ND)	Record Qualifier	Used to indicate exam not done. Should be null if a result exists in BSORRES.	Perm
BSREASND	Reason Test Not Done	Char		Record Qualifier	Describes why a measurement or test was not performed such as BROKEN EQUIPMENT, SPECIMEN CONDITION or SPECIMEN LOST. Used in conjunction with BSSTAT when value is NOT DONE.	Perm
BSNAM	Vendor Name	Char		Record Qualifier	The name or identifier of the laboratory that performed the test.	Perm
BSSPEC	Specimen Type	Char	*	Record Qualifier	Defines the type of specimen used for a measurement. Examples: SERUM, PLASMA, URINE, SOFT TISSUE.	Perm
BSPARENT	Specimen Parent	Char		Record Qualifier	Identifies the parent of a specimen to support tracking its genealogy.	Perm
BSSPCLVL	Specimen Level	Num		Record Qualifier	Identifies the generation number of the sample where the collected sample is considered the first generation.	Perm
BSANTREG	Anatomical Region	Char	*	Variable Qualifier of BSSPEC	Defines the specific anatomical or biological region of a tissue, organ specimen or the region from which the specimen is obtained, as defined in the protocol, such as a section or part of what is described in the --SPEC variable. Examples: CORTEX, MEDULLA, MUCOSA, CEREBRAL AQUEDECT	Perm
BSSPCCND	Specimen Condition	Char	*	Record Qualifier	Free or standardized text describing the condition of the specimen.Examples: HEMOLYZED, ICTERIC, LIPEMIC.	Perm
BSMETHOD	Method of Test or Examination	Char	*	Record Qualifier	Method of the test or examination. Example: EXCISION	Perm
BSRUNID	Run ID	Char		Record Qualifier	A unique identifier for a particular run of a test on a particular batch of samples.	Perm

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Role	CDISC Notes	Core
BSBLFL	Baseline Flag	Char	(NY)	Record Qualifier	Indicator used to identify a baseline value.	Exp
VISITNUM	Visit Number	Num		Timing	1. Clinical encounter number. 2. Numeric version of VISIT, used for sorting.	Exp
VISIT	Visit Name	Char		Timing	1. Protocol-defined description of clinical encounters. 2. May be used in addition to VISITNUM and/or VISITDY	Perm
VISITDY	Planned Study Day of Visit	Num		Timing	Planned study day of the visit based upon RFSTDTC in Demographics.	Perm
BSDTC	Date/Time of Specimen Collection	Char	ISO 8601	Timing	Date/time of specimen collection	Exp
BSSTDTC	Start Date/Time of Specimen Process	Char	ISO 8601	Timing	Start Date/time of specimen handling process	Perm
BSENDTC	End Date/Time of Specimen Process	Char	ISO 8601	Timing	End Date/time of specimen handling process	Perm
BSDY	Study Day of Specimen Collection	Num		Timing	1. Study day of specimen collection, measured as integer days. 2. Algorithm for calculations must be relative to the sponsor-defined RFSTDTC variable in Demographics. This formula should be consistent across the submission.	Perm
BSTPT	Planned Time Point Name	Char		Timing	1. Text Description of time when specimen should be taken. 2. This may be represented as an elapsed time relative to a fixed reference point, such as time of last dose. See BSTPTNUM and BSTPTREF. Examples**:** Start, 5 min post.	Perm
BSTPTNUM	Planned Time Point Number	Num		Timing	Numerical version of BSTPT to aid in sorting.	Perm
BSELTM	Planned Elapsed Time from Time Point Ref	Char	ISO 8601	Timing	Planned Elapsed time (in ISO 8601) relative to a planned fixed reference (BSTPTREF). This variable is useful where there are repetitive measures. Not a clock time or a date time variable. Represented as an ISO 8601 duration. Examples**:** '-PT15M' to represent the period of 15 minutes prior to the reference point indicated by BSTPTREF, or 'PT8H' to represent the period of 8 hours after the reference point indicated by BSTPTREF.	Perm
BSTPTREF	Time Point Reference	Char		Timing	Name of the fixed reference point referred to by BSELTM, if used for BSTPTNUM, and BSTPT. Examples**:** PREVIOUS DOSE, PREVIOUS MEAL.	Perm
BSRFTDTC		Char	ISO 8601	Timing	Date/time of the reference time point, BSTPTREF.	Perm

** Indicates variable may be subject to controlled terminology, (Parenthesis indicates CDISC/NCI codelist code value)*

3.2.3 Assumptions for Biospecimen Domain Model

1. The BS domain is used to store findings related to specimen handling. It also contains specimen characteristics such as type, amount, or size.

3.3 Examples for Biospecimen Events and Biospecimen Domain Models

Example 1

Cell-free RNA, which can be obtained from plasma, may be useful for some tumor-specific cancer detection[1], but has poor integrity[2]. In this example, a blood sample was drawn, centrifuged to get plasma, and treated with EDTA, and then the RNA was extracted and purified. The SPDEVID for the RNA purification kit is given in row 4.

be.xpt

Row	STUDYID	USUBJID	BESPDEVID	BESEQ	BEREFID	BETERM	BEDECOD	BECAT	BESCAT	VISITNUM	BEDTC	BESTDTC	BEENDTC
1	3441271	MU-298		1	298B1	Collected	COLLECTED	COLLECTION	BLOOD	2	2010-04-01		
2	3441271	MU-298		2	298P1	Extraction	EXTRACTED	EXTRACTION	PLASMA				
						Treated	TREATED	PREP	PLASMA				
3	3441271	MU-298		3	298R1	Genetic Extraction	GENETIC EXTRACTION	EXTRACTION	RNA	2	2010-04-01		
4	3441271	MU-298	PURKIT	4	298R1	Purified	PURIFIED	PREP					

suppbe.xpt

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
1	3441271	BE	AAA-XXX1	BETERM	Treated	TRT	Treatment	EDTA

Example 2

This is an example of volumetric measurements for specimens.

bs.xpt

Row	STUDYID	DOMAIN	USUBJID	BSSEQ	BSREFID	BSTESTCD	BSTEST	BSCAT	BSORRES	BSORRESU	BSSTRESC	BSSTRESN	BSSTRESU	BSBLFL	VISITNUM	BSDTC
1	ISI1414	BS	XER01	1	1	VOLUME	Volume		10	mL	10	10	mL		3	2012-12-03
2	ISI1414	BS	XER01	2	2	VOLUME	Volume		4	mL	4	4	mL		3	2012-12-03
3	ISI1414	BS	XER01	3	1-1	VOLUME	Volume		3	mL	3	3	mL		3	2012-12-03
4	ISI1414	BS	XER01	4	1-2	VOLUME	Volume		3	mL	3	3	mL		3	2012-12-03

Example 3

This example shows data about RNA integrity.

bs.xpt

Row	STUDYID	DOMAIN	USUBJID	BSSEQ	BSREFID	BSTESTCD	BSTEST	BSCAT	BSORRES	BSSTRESC	BSSTRESN	BSXFN
1	A12345	BS	43871	1	1148.26704	A260A230	A260/A230	QC	2.05	2.05	2.05	2.16.090.1.135764.3.4:7280912
2	A12345	BS	43871	2	1148.26704	A260A280	A260/A280	QC	2	2	2	2.16.090.1.135764.3.4:7280912
3	A12345	BS	43871	3	1148.26704	I28S18S	28S/18S	QC	1.2	1.2	1.2	2.16.090.1.135764.3.4:7280912
4	A12345	BS	43871	4	1148.26704	RIN	RNA INTEGRITY NUMBER	QC	9.5	9.5	9.5	2.16.090.1.135764.3.4:7280912
5	A12345	BS	43871	5	1148.26704	CLSSCORE	Classifier Score	QC	0.4	0.4	0.4	2.16.090.1.135764.3.4:7280912

[1] Tsui NB, Ng EK, Lo YM. Molecular analysis of circulating RNA in plasma. *Methods Mol Biol.* 2006;336:123-34.
[2] Cerkovnik P, Perhavec A, Zgajnar J, Novakovic S. Optimization of an RNA isolation procedure from plasma samples. *Int J Mol Med.* 2007;20(3):293-300.

Row	BSNAM	BSMETHOD	BSRUNID	VISIT	VISITNUM	VISITDY	BSDTC
1 (cont)	Deluxe Central Labs	SPECTROPHOTOMETRY	1000450001	Baseline	1	1	2005-03-21T11:28:17
2 (cont)	Deluxe Central Labs	SPECTROPHOTOMETRY	1000450001	Baseline	1	1	2005-03-21T11:28:17
3 (cont)	Deluxe Central Labs	ELECTROPHORESIS	1000450001	Baseline	1	1	2005-03-21T11:28:17
4 (cont)	Deluxe Central Labs	ELECTROPHORESIS	1000450001	Baseline	1	1	2005-03-21T11:28:17
5 (cont)	Deluxe Central Labs	BAYESIAN	1000450001	Baseline	1	1	2005-03-21T11:28:17

Example 4
In this example, a specimen is collected, flash frozen, thawed, and shipped to another location.

Some tests are very sensitive to specimen handling processes such as flash freezing or time spent in transit. Therefore, it is important to record when the processes were started and completed. Such information goes is placed in the BE domain.

Row 1: Shows specimen collection.
Rows 2-4: Show the start and end date/times of flash freezing, storing while frozen, and thawing.
Row 5: Records the transportation of a biospecimen. The point of origination is not given because it is assumed to be the study site.

be.xpt

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	BESEQ	BEREFID	BETERM	BEDECOD	BEPARTY	BEPRTYID	BECAT	BESCAT
1	ABC134	BE	43871	TS409871	1	1148.267	Excision	EXCISION			COLLECTION	SOFT TISSUE
2	ABC134	BE	43871		2	1148.267	Flash Frozen	FLASH FROZEN			PREP	
3	ABC134	BE	43871	309827	3	1148.267	Stored in Freezer	STORED			STORING	
4	ABC134	BE	43871		4	1148.267	Thaw	THAW			PREP	
5	ABC134	BE	43871	LN43871	5	1148.267	Shipped	SHIPPED	ABC LAB	01	TRANSPORT	

Row	BODSYS	BELOC	VISITNUM	VISIT	BEDTC	BESTDTC	BEENDTC
1 (cont)	Nervous System [A08]	BRAIN	1	BASELINE	2005-02-20	2005-03-20T15:07	
2 (cont)	Nervous System [A08]	BRAIN	1	BASELINE	2005-03-20	2005-03-20T15:07	2005-03-20T13:22
3 (cont)	Nervous System [A08]	BRAIN	1	BASELINE	2005-03-20	2005-03-20T13:22	2005-03-21T10:29
4 (cont)	Nervous System [A08]	BRAIN	1	BASELINE	2005-03-21	2005-03-21T10:29	2005-03-21T10:36
5 (cont)	Nervous System [A08]	BRAIN	1	BASELINE	2005-03-21	2005-03-21T11:00	2005-03-21T15:00

Findings related to specimen handling processes are stored in BS. These processes are important to maintain the integrity of the specimens used in genetic variation and gene expression testing. Depending on how a study is designed, there might be very specific specimen handling specifications contained in the protocol for all labs to follow. Other protocols may let the labs determine the processes to follow. The examples below support the latter approach.

Row 1: Contains the volume of the biospecimen.
Row 2: Depicts the temperature to which the specimen was flash-frozen.

bs.xpt

Row	STUDYID	DOMAIN	USUBJID	BSSEQ	BSREFID	BSTESTCD	BSTEST	BSCAT	BSORRES	BSORRESU	BSSTRESC	BSSTRESN
1	ABC134	BS	43871	1	1148.267	VOLUME	Volume	SPECIMEN MEASUREMENT	2	cm3	2	2

Row	STUDYID	DOMAIN	USUBJID	BSSEQ	BSREFID	BSTESTCD	BSTEST	BSCAT	BSORRES	BSORRESU	BSSTRESC	BSSTRESN
2	ABC134	BS	43871	2	1148.267	FFRZTMP	Flash Frozen Temp	SPECIMEN HANDLING	-80	C	-80	

Row	BSSTRESU	BSPEC	BSANTREG	BSBLFL	VISITNUM	BSDTC
1 (cont)	cm3	BRAIN	CEREBRAL AQUEDECT			
2 (cont)	C	BRAIN	CEREBRAL AQUEDECT	Y	1	2005-03-20

RELREC relates the records in BE and BS to each other.

Rows 1-2: Tie the specimen volume to its collection
Rows 3-4: Tie the temperature at which the specimen was flash frozen to the event of its occurrence.

relrec.xpt

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	RELTYPE	RELID
1	ABC134	BE	43871	BESEQ	1		1
2	ABC134	BS	43871	BSSEQ	1		1
3	ABC134	BE	43871	BESEQ	2		2
4	ABC134	BS	43871	BSSEQ	2		2

3.4 Related Specimens

There has been some question as to how to record specimen relationships and hierarchies, such as when a specimen is resectioned or aliquoted, and it is important to keep track of the chain of parent-child relationships back to the original collected specimen. One possibility would be to add two variables, --PARENT, and --SPCLVL to the BE domain, or to the BS domain, or to both. Another method would be to create a special dataset specifically for the purpose of holding specimen relationship data, as given below.

3.4.1 RELSPEC Dataset

relrec.xpt, Related Specimens, Version 3.x.x. One record per specimen.

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Role	CDISC Notes	Core
STUDYID	Study Identifier	Char		Identifier	Unique identifier for a study.	Req
USUBJID	Unique Subject Identifier	Char		Identifier	Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product.	Req
REFID	Specimen ID	Char		Identifier	Specimen identifier, unique within USUBJID.	Req
SPEC	Specimen Type	Char	*	Record Qualifier	Defines the type of specimen used for a measurement. Examples**:** SERUM, PLASMA, URINE, SOFT TISSUE.	Perm
PARENT	Specimen Parent	Char		Record Qualifier	Identifies the parent of a specimen to support tracking its genealogy.	Exp

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Role	CDISC Notes	Core
SPCLVL	Specimen Level	Num		Record Qualifier	Identifies the generation number of the sample where the collected sample is considered the first generation.	Req

** Indicates variable may be subject to controlled terminology, (Parenthesis indicates CDISC/NCI codelist code value)*

3.4.2 Examples for Related Specimens

The following examples all use the same sample specimen geneology, as illustrated below.

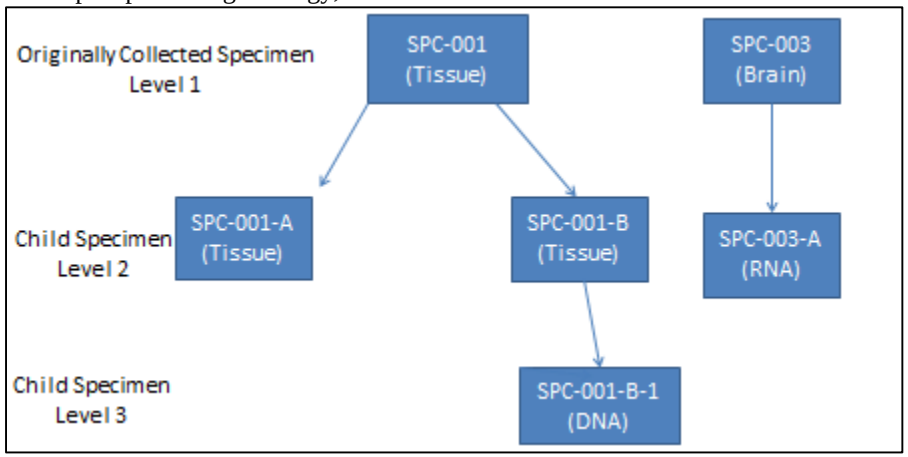


Figure 1: Sample Specimen Relationship

Any specimen with a BSSPCLVL value of "1" and a blank value for BSPARENT indicates a "collected sample". All other values represent a "derived sample". BSSPEC will always reflect the specimen type for the collected sample. BSGENSMP will be populated only for derived genetic samples.

Example 1
This example shows how the specimen genealogy might be represented in the BE domain.

Row 1: Specimen SPC-001 is the collected sample.
Rows 2-3: Specimen SPC-001 underwent a process to produce two derived samples "SPC-001-A" and "SPC-001-B". BSPARENT indicates that "SPC-001" is the immediate parent. Both derived samples contain a BSSPCLVL of "2" to indicate this is a second generation sample set derived from the collected sample.
Row 4: Specimen SPC-001-B is a DNA sample that was extracted from the derived sample SPC-001-, therefore it is considered a third generation (BSPSCLVL="3").
Row 5: This represents a second collected specimen.
Row 6: This RNA was extracted directly from the collected specimen so it has a second generation value of "2" in BSSPCLVL.

be.xpt

Row	STUDYID	DOMAIN	USUBJID	BESEQ	BEREFID	BETERM	BEDECOD	BECAT	BESCAT	BEPARENT	BESPCLVL	…
1	ABC-123	BS	001-01	1	SPC-001	Excision	EXCISION	COLLECTION	TISSUE		1	…
2	ABC-123	BS	001-01	2	SPC-001-A	Resectioning	RESECTIONING	PREP	TISSUE	SPC-001	2	…
3	ABC-123	BS	001-01	3	SPC-001-B	Resectioning	RESECTIONING	PREP	TISSUE	SPC-001	2	…
4	ABC-123	BS	001-01	4	SPC-001-B-1	Genetic Extraction	GENETIC EXTRACTION	EXTRACTION	DNA	SPC-001-B	3	…
5	ABC-123	BS	002-01	5	SPC-003	Excision	EXCISION	COLLECTION	BRAIN		1	…
6	ABC-123	BS	002-01	6	SPC-003-A	Genetic Extraction	GENETIC EXTRACTION	EXTRACTION	RNA	SPC-003	2	…

Example 2

This example shows how the specimen genealogy might be represented in the BS domain.

Row 1: Specimen SPC-001 is the collected samples as indicated via BSSPCLVL = "1".

Rows 2-3: Specimen SPC-001 underwent a process to produce two derived samples "SPC-001-A" and "SPC-001-B". BSPARENT indicates that "SPC-001" is the immediate parent. Both derived samples contain a BSSPCLVL of "2" to indicate this is a second generation sample set derived from the collected sample.

Row 4: Specimen SPC-001-B-1 is a DNA sample that was extracted from the derived sample SPC-001-B, therefore it is considered a third generation (BSPSCLVL="3").

Row 5: This represents a second collected specimen. We used "BRAIN" as the specimen type since this is a valid specimen for toxicology studies as commonly used by SEND.

Row 6: This RNA was extracted directly from the collected specimen so it has a second generation value of "2" in BSSPCLVL.

bs.xpt

Row	STUDYID	DOMAIN	USUBJID	BSSEQ	BSREFID	BSTESTCD	BSTEST	BSORRES	BSORRESU	…	BSSPEC	BSPARENT	BSSPCLVL	…
1	ABC-123	BS	001-01	1	SPC-001	SPECVOL	Biospecimen Volume	100			TISSUE		1	…
2	ABC-123	BS	001-01	2	SPC-001-A	SPECVOL	Biospecimen Volume	50			TISSUE	SPC-001	2	…
3	ABC-123	BS	001-01	3	SPC-001-B	SPECVOL	Biospecimen Volume	50			TISSUE	SPC-001	2	…
4	ABC-123	BS	001-01	4	SPC-001-B-1	SPECVOL	Biospecimen Volume				DNA	SPC-001-B	3	…
5	ABC-123	BS	002-01	5	SPC-003	SPECVOL	Biospecimen Volume	100			BRAIN		1	…
6	ABC-123	BS	002-01	6	SPC-003-A	SPECVOL	Biospecimen Volume				RNA	SPC-003	2	…

Example 3

This example shows how the specimen genealogy might be represented in the RELSPEC dataset.

Row	STUDYID	USUBJID	REFID	SPEC	PARENT	SPCLVL
1	ABC-123	001-01	SPC-001	TISSUE		1
2	ABC-123	001-01	SPC-001-A	TISSUE	SPC-001	2
3	ABC-123	001-01	SPC-001-B	TISSUE	SPC-001	2
4	ABC-123	001-01	SPC-001-B-1	DNA	SPC-001-B	3
5	ABC-123	001-01	SPC-003	BRAIN		1
6	ABC-123	001-01	SPC-003-A	RNA	SPC-003	2

6 Domain Models Based on the General Observation Classes

6.3 Findings

Nervous System Findings (NV)

NV - Description/Overview for Nervous System Findings Domain Model

A domain for physiological findings related to the nervous system, including the brain, spinal cord, the cranial and spinal nerves, autonomic ganglia and plexuses.

The NV domain is still in the very early stages of development.

NV - Specification for Nervous System Findings Domain Model

nv.xpt, Nervous System Findings — Findings, Version 3.3. One record per Nervous System Findings finding per location per time point per visit per subject, Tabulation

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Role	CDISC Notes	Core
STUDYID	Study Identifier	Char		Identifier	Unique identifier for a study.	Req
DOMAIN	Domain Abbreviation	Char	NV	Identifier	Two-character abbreviation for the domain.	Req
USUBJID	Unique Subject Identifier	Char		Identifier	Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product.	Req
NVSEQ	Sequence Number	Num		Identifier	Sequence Number given to ensure uniqueness of subject records within a domain. May be any valid number.	Req
NVGRPID	Group ID	Char		Identifier	Used to tie together a block of related records in a single domain for a subject.	Perm
NVREFID	Reference ID	Char		Identifier	Internal or external procedure identifier.	Perm
NVSPID	Sponsor-Defined Identifier	Char		Identifier	Sponsor-defined reference number. Perhaps pre-printed on the CRF as an explicit line identifier or defined in the sponsor's operational database. Example: Line number from the procedure or test page.	Perm
NVLNKID	Link ID	Char		Identifier	Identifier used to link a procedure to the assessment results over the course of the study.	Perm

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Role	CDISC Notes	Core
NVTESTCD	Test or Examination Short Name	Char	*	Topic	Short name of the measurement, test, or examination described in NVTEST. It can be used as a column name when converting a dataset from a vertical to a horizontal format. The value in NVTESTCD cannot be longer than 8 characters, nor can it start with a number (e.g., "1TEST"). NVTESTCD cannot contain characters other than letters, numbers, or underscores. Example: SUVR	Req
NVTEST	Test or Examination Name	Char	*	Synonym Qualifier	Verbatim name of the test or examination used to obtain the measurement or finding. The value in NVTEST cannot be longer than 40 characters. Example: Standard Uptake Value Ratio, etc.	Req
NVCAT	Category for Test	Char	*	Grouping Qualifier	Used to categorize observations across subjects.	Perm
NVSCAT	Subcategory for Test	Char	*	Grouping Qualifier	A further categorization.	Perm
NVPOS	Position of Subject	Char	(POSITION)	Record Qualifier	Position of the subject during a measurement or examination. Examples: SUPINE, STANDING, SITTING.	Perm
NVORRES	Result or Finding in Original Units	Char		Result Qualifier	Result of the procedure measurement or finding as originally received or collected.	Exp
NVORRESU	Original Units	Char	(UNIT)	Variable Qualifier	Original units in which the data were collected. The unit for NVORRES.	Perm
NVSTRESC	Character Result/Finding in Std Format	Char		Result Qualifier	Contains the result value for all findings, copied or derived from NVORRES in a standard format or standard units. NVSTRESC should store all results or findings in character format; if results are numeric, they should also be stored in numeric format in NVSTRESN.	Exp
NVSTRESN	Numeric Result/Finding in Standard Units	Num		Result Qualifier	Used for continuous or numeric results or findings in standard format; copied in numeric format from NVSTRESC. NVSTRESN should store all numeric test results or findings.	Perm
NVSTRESU	Standard Units	Char	(UNIT)	Variable Qualifier	Standardized unit used for NVSTRESC or NVSTRESN.	Perm
NVSTAT	Completion Status	Char	(ND)	Record Qualifier	Used to indicate a test was not done, or a measurement was not taken. Should be null if a result exists in NVORRES.	Perm
NVREASND	Reason Test Not Performed	Char		Record Qualifier	Describes why a measurement or test was not performed. Examples: BROKEN EQUIPMENT or SUBJECT REFUSED. Used in conjunction with NVSTAT when value is NOT DONE.	Perm
NVXFN	Raw Data File	Char		Record Qualifier	Filename for an external file used to populate the NV domain.	Perm
NVNAM	Laboratory/Vendor Name	Char		Record Qualifier	Name or identifier of the vendor (e.g., laboratory) that provided the test results.	Perm

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Role	CDISC Notes	Core
NVLOC	Location Used for Measurement	Char	(LOC)	Record Qualifier		Perm
NVLAT	Specimen Laterality within Subject	Char	(LAT)	Variable Qualifier	Qualifier for laterality of the specimen within the subject for paired specimens. Examples: LEFT, RIGHT, BILATERAL.	Perm
NVDIR	Specimen Directionality within Subject	Char	(DIR)	Variable Qualifier	Qualifier for directionality of the specimen within the subject. Examples: DORSAL, PROXIMAL.	Perm
NVMETHOD	Method of Procedure Test	Char	(METHOD)	Record Qualifier	Method of the procedure.	Perm
NVBLFL	Baseline Flag	Char	(NY)	Record Qualifier	Indicator used to identify a baseline value. The value should be "Y" or null.	Exp
NVDRVFL	Derived Flag	Char	(NY)	Record Qualifier	Used to indicate a derived record. The value should be Y or null. Records which represent the average of other records, or that do not come from the CRF, or are not as originally collected or received are examples of records that would be derived for the submission datasets. If NVDRVFL=Y, then NVORRES would be null, with, and (if numeric) NVSTRESN having the derived value.	Perm
NVEVAL	Evaluator	Char	*	Record Qualifier	Role of the person who provided the evaluation. Used only for results that are subjective (e.g., assigned by a person or a group). Should be null for records that contain collected or derived data. Examples: INVESTIGATOR, ADJUDICATION COMMITTEE, VENDOR.	Perm
VISITNUM	Visit Number	Num		Timing	1. Clinical encounter number. 2. Numeric version of VISIT, used for sorting.	Exp
VISIT	Visit Name	Char		Timing	1. Protocol-defined description of clinical encounter. 2. May be used in addition to VISITNUM and/or VISITDY.	Perm
VISITDY	Planned Study Day of Visit	Num		Timing	Planned study day of the visit based upon RFSTDTC in Demographics.	Perm
NVDTC	Date/Time of Test	Char	ISO 8601	Timing	Date of procedure or test.	Exp
NVDY	Study Day of Test	Num		Timing	1. Study day of the procedure or test, measured as integer days. 2. Algorithm for calculations must be relative to the sponsor-defined RFSTDTC variable in Demographics.	Perm
NVTPT	Planned Time Point Name	Char		Timing	1. Text Description of time when measurement should be taken. 2. This may be represented as an elapsed time relative to a fixed reference point, such as time of last dose. See NVTPTNUM and NVTPTREF. Examples: Start, 5 min post.	Perm
NVTPTNUM	Planned Time Point Number	Num		Timing	Numerical version of NVTPT to aid in sorting.	Perm
NVELTM	Planned Elapsed Time from Time Point Ref	Char	ISO 8601	Timing	Planned elapsed time (in ISO 8601) relative to a fixed time point reference (NVTPTREF). Not a clock time or a date time variable. Represented as an ISO 8601 duration. Examples: "-PT15M" to represent the period of 15 minutes prior to the reference point indicated by NVTPTREF, or "PT8H" to represent the period of 8 hours after the reference point indicated by NVTPTREF.	Perm

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Role	CDISC Notes	Core
NVTPTREF	Time Point Reference	Char		Timing	Name of the fixed reference point referred to by NVELTM, NVTPTNUM, and NVTPT. Examples: PREVIOUS DOSE, PREVIOUS MEAL.	Perm
NVRFTDTC	Date/Time of Reference Time Point	Char	ISO 8601	Timing	Date/time of the reference time point, NVTPTREF.	Perm

** Indicates variable may be subject to controlled terminology (Parenthesis indicates CDISC/NCI codelist code value)*

NV - Assumptions for Nervous System Findings Domain Model

1. NV Definition: This domain has been designed to store data on neurological physiological findings that include information relating to the brain, the spinal cord, and the nerves.
2. Additional Findings Qualifiers
    a. The following variables would not generally be used in NV: --LOINC, --FAST, --TOX, --TOXGR.

NV - Examples for Nervous System Findings Domain Model

Example 1
This example shows measures for standard uptake value ratios taken from three PET scans. NVDTC corresponds to the date of the PET or PET/CT procedure from which these results were obtained.

Rows 1-2: Show the Standard Uptake Value Ratio (SUVR) findings based on a PET/CT scan for subject AD01-101.
Rows 3-4: Show the SUVR findings based on a PET/CT scan for subject AD01-102.
Rows 5-6: Show the SUVR findings based on an FDG-PET scan for subject AD AD01-103.

nv.xpt

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	NVSEQ	NVREFID	NVLNKID	NVTESTCD	NVTEST	NVORRES	NVORRESU
1	ABC123	NV	AD01-101	22	1	1236	03	SUVR	Standard Uptake Value Ratio	.95	RATIO
2	ABC123	NV	AD01-101	22	2	1236	03	SUVR	Standard Uptake Value Ratio	1.17	RATIO
3	ABC123	NV	AD01-102	22	1	1237	04	SUVR	Standard Uptake Value Ratio	1.21	RATIO
4	ABC123	NV	AD01-102	22	2	1237	04	SUVR	Standard Uptake Value Ratio	1.78	RATIO
5	ABC123	NV	AD01-103	44	1	1238	05	SUVR	Standard Uptake Value Ratio	1.52	RATIO
6	ABC123	NV	AD01-103	44	2	1238	05	SUVR	Standard Uptake Value Ratio	1.63	RATIO

Row	NVSTRESC	NVSTRESN	NVSTRRESU	NVLOC	NVDIR	NVMETHOD	NVDTC
1 (cont)	.95	.95	RATIO	PRECUNEUS		PET/CT SCAN	2012-05-22
2 (cont)	1.17	1.17	RATIO	CINGULATE CORTEX	POSTERIOR	PET/CT SCAN	2012-05-22
3 (cont)	1.21	1.21	RATIO	PRECUNEUS		PET/CT SCAN	2012-05-22
4 (cont)	1.78	1.78	RATIO	CINGULATE CORTEX	POSTERIOR	PET/CT SCAN	2012-05-22
5 (cont)	1.52	1.52	RATIO	PRECUNEUS		FDGPET	2012-05-22
6 (cont)	1.63	1.63	RATIO	CINGULATE CORTEX	POSTERIOR	FDGPET	2012-05-22

A supplemental qualifiers dataset is used for additional data elements that are not part of the NV domain.

Rows 1-6: Shows the reference region used for the SUVR tests shown in the NV domain.

suppnv.xpt

Row	STUDYID	RDOMAIN	USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
1	ABC123	NV	AD01-101	NVSEQ	1	REFREG	Reference Region	CEREBELLUM
2	ABC123	NV	AD01-101	NVSEQ	2	REFREG	Reference Region	CEREBELLUM
3	ABC123	NV	AD01-102	NVSEQ	1	REFREG	Reference Region	CEREBELLUM
4	ABC123	NV	AD01-102	NVSEQ	2	REFREG	Reference Region	CEREBELLUM
5	ABC123	NV	AD01-103	NVSEQ	1	REFREG	Reference Region	PONS
6	ABC123	NV	AD01-103	NVSEQ	2	REFREG	Reference Region	PONS

4 Pharmacogenomics/Genetics Domains

4.1 Pharmacogenomics Findings (PF)

The Pharmacogenomics Findings (PF) domain is a findings domain for gene expression and genetic variation assessments.

--TESTCV, --TCVNM, and --TSCVVR comprise a set of variables that propose to hold external terminology for a given --TEST, so that individual variables need not be created for each possible dictionary or database (such as the --LOINC variable). --RESCV, --RCVNM, and RCVVR comprise a comparable set of variables for results.

PFSPCIES and PFSTRAIN are proposed variables for recording the species and strain of a subject's pathogen (such as a virus) when that pathogen comes under study.

PFGENTYP, PFGENRI, and PFGENSRI are variables used to further identify the specific portion of the genome being tested. For genetic variation, PFREFSEQ, PFGENLOC, PFORRES and PFORREF hold individual aspects of the variation recorded, which is given in HGVS nomenclature in PFSTRESC, while PFMUTYP is used to record whether the variation is inheritable or arising only in parts of the body such as in cancerous tumors.

4.1.1 Specification for Pharmacogenomics Findings Domain Model

pf.xpt, Pharmacogenomics Findings - Findings. One record per method/setup observation per specimen collected per date of test per subject, Tabulation

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Role	CDISC Notes	Core
STUDYID	Study Identifier	Char		Identifier	Unique identifier for a study within the submission.	Req
DOMAIN	Domain Abbreviation	Char	PF	Identifier	Two-character abbreviation for the domain most relevant to the observation.	Req
USUBJID	Unique Subject Identifier	Char		Identifier	Unique subject identifier within the submission.	Req
SPDEVID	Sponsor Device Identifier	Char		Identifier	Sponsor-defined identifier for a device.	Perm
PFSEQ	Sequence Number	Num		Identifier	Sequence number given to ensure uniqueness within a dataset for a subject. Can be used to join related records.	Req
PFGRPID	Group ID	Char		Identifier	Used to tie together a block of related records in a single domain to support relationships within the domain and between domains. Example**:** drug metabolism link to PK data domain or grouping results related to one test; NS5B_71 where NS5B is the mutation and 71 is the position of the codon or just a numeric assign to group the records.	Perm
PFREFID	Reference ID	Char		Identifier	Optional internal or external identifier such as lab specimen ID.	Perm
PFSPID	Sponsor ID	Char		Identifier	Sponsor-defined identifier.	Perm
PFLNKID	Link ID	Char		Identifier	Supports linking information across different domains.	Perm

PFTESTCD	Genomics Test Short Name	Char	*	Topic	Short name for the test. Examples: AA, NUC, ALE, GENOTYPE, PVAL, FOLDCHG.	Req
PFTEST	Genomics Test Name	Char	*	Synonym Qualifier	The verbatim name used to obtain the measurement or finding. Examples: Amino Acid, Nucleotide, Allele, Genotype, P Value, Fold Change.	Perm
PFTESTCV	Genomics Test Controlled Vocabulary	Char	*	Synonym Qualifier	A code or identifier for the test. The length of this variable can be longer than 8 to accommodate the length of the external terminology. Example: 48019-4.	Perm
PFTCVNM	Name of the Vocab for the Test	Char		Result Qualifier	The name of the terminology from which PFTESTCV is taken. Examples: SNOMED, LOINC.	Perm
PFTCVVR	Version of the Vocab for the Test	Char		Result Qualifier	The version of the terminology, if applicable.	Perm
PFGENTYP	Genetic Region of Interest Type	Char	*	Result Qualifier	Identifies the type of genetic region of interest. Examples: GENE, SECTOR, PROTEIN.	Exp
PFGENRI	Genetic Region of Interest	Char	*	Result Qualifier	Area within the DNA sequences. Example: EGFR, CIC, KRAS	Exp
PFGENSRI	Genetic Sub-Region of Interest	Char	*	Result Qualifier	Area within the DNA sequences that is contained within the region identified by PFGENRI. Example: PROTEASE	Perm
PFSPCIES	Biological Classification	Char	*	Grouping Qualifier	Biological classifications for an organism capable of breeding and producing offspring. Examples: BACTERIUM, HCV, HIV	Perm
PFSTRAIN	Type of Strain	Char	*	Grouping Qualifier	A genetic variant or subtype of a microorganism. Examples: 1a, 1b.	Perm
PFCAT	Category for Pharmacogenomics Lab Test	Char	*	Grouping Qualifier	Used to categorize types of genetic/genomic tests. Examples: GENETIC VARIATION, GENE EXPRESSION	Exp
PFSCAT	Subcategory for Pharmacogenomics Lab Test	Char	*	Grouping Qualifier	A further categorization of the various test types based on particular characteristics of a test. Examples: INTERPRETATION, PHENOTYPIC EXPRESSION	Perm
PFGENLOC	Genetic Location	Char		Result Qualifier	Specifies a location within a sequence pertaining to the observed results contained in PFORRES, PFSTRESC and PFSTRESN.	Perm
PFORRES	Result or Finding in Original Units	Char		Result Qualifier	Result of the measurement or finding as originally received or collected. Examples: T, Gly, del.	Exp
PFORRESU	Original Units	Char		Variable Qualifier	Represents the unit of measure used by PFORRES, if applicable. Examples: copies/5uL, LOG10 IU/ml, Cycles	Perm
PFORREF	Reference Result	Char		Variable Qualifier	Reference result for the measurement or finding, in the same format as PFORRES. PFORREF uses the same units as PFORRES, if applicable.	Perm
PFSTRESC	Result or Finding in Standard Format	Char		Result Qualifier	Contains the result value for all findings, copied or derived from PFORRES in a standard format or in standard units. PFSTRESC should store all results or findings in character format; if results are numeric, they should also be stored in numeric format in PFSTRESN.	Exp
PFSTRESN	Numeric Result/Finding in Standard Units	Num		Result Qualifier	Used for continuous or numeric results or findings in standard format; copied in numeric format from PFSTRESC. PFSTRESN should store all numeric test results or findings. Example: for P-Value; 0.5391	Perm
PFSTRESU	Standard Units	Char	(UNIT)	Variable Qualifier	Standardized units used for PFSTRESC and PFSTRESN.	Perm
PFREFSEQ	Reference Sequence	Char	*	Variable Qualifier	A unique identifier for the reference sequence used to identify the genetic variation. Examples: NM_001234, rs02973492	Perm

PFRESCV	Genomics Result Controlled Vocabulary	Char	*	Result Qualifier	A code or identifier for the result in PFSTRESC. The length of this variable can be longer than 8 to accommodate the length of the external terminology. Examples: Arg, C49488, rs1042522.	Perm
PFRCVNM	Name of the Vocab for the Result	Char		Result Qualifier	The name of the terminology or database from which PFRESCV is taken. Examples: HGVS, dbSNP	Perm
PFRCVVR	Version of the Vocab for the Result	Char		Result Qualifier	The version of the terminology, if applicable.	Perm
PFRESCAT	Result Category	Char	*	Result Qualifier	Used to categorize the result of a finding. Example: RESISTANCE VARIANT.	Perm
PFMUTYP	Mutation Type	Char	*	Grouping Qualifier	Indicates whether a mutation is inheritable or not. Examples: SOMATIC, GERMLINE	Perm
PFSTAT	Test Status	Char	(ND)	Record Qualifier	Used to indicate exam not done. Should be null if a result exists in PFSTRESC.	Perm
PFREASND	Reason Test Not Done	Char		Record Qualifier	Describes why a measurement or test was not performed such as BROKEN EQUIPMENT, SUBJECT REFUSED, or SPECIMEN LOST. Used in conjunction with PFSTAT when value is NOT DONE.	Perm
PFXFN	Raw Data File or Life Science Identifier	Char		Record Qualifier	Direct reference identifier for Microarray or Genotypic data contained in a separate file in its native format.	Perm
PFNAM	Vendor Name	Char		Record Qualifier	Name or identifier of the laboratory or biotech firm who provides the test results.	Perm
PFSPEC	Specimen Type	Char	*	Record Qualifier	Defines the type of specimen used for a measurement. Examples: DNA, RNA	Perm
PFSPCCND	Specimen Condition	Char		Record Qualifier	Free or standardized text describing the condition of the specimen. Example: CONTAMINATED	Perm
PFMETHOD	Method Code for Test	Char	(METHOD)	Record Qualifier	Special instructions for the execution of genomics or genetic testing. Examples: SNPASSAY, CLIP SEQUENCING, PYROSEQUENCING, BICHROME GENE EXPRESSION CHIP).	Req
PFRUNID	Run ID	Char		Record Qualifier	A unique identifier for a particular run of a test on a particular batch of samples.	Perm
PFANMETH	Analysis Method	Char	*	Record Qualifier	Analysis method applied to obtain a summarized result. Analysis method describes the method of secondary processing applied to a complex observation result (e.g. an image or a genetic sequence).	Perm
PFBLFL	Baseline Flag	Char	(NY)	Record Qualifier	Indicator used to identify a baseline value.	Exp
PFDRVFL	Derived Flag	Char	(NY)	Record Qualifier	Used to indicate a derived record.	Perm
VISITNUM	Visit Number	Num		Timing	1. Clinical encounter number. 2. Numeric version of VISIT, used for sorting.	Exp
VISIT	Visit Name	Char		Timing	1. Protocol-defined description of clinical encounter. 2. May be used in addition to VISITNUM and/or VISITDY	Perm
VISITDY	Planned Study Day of Visit	Num		Timing	Planned study day of the visit based upon RFSTDTC in Demographics.	Perm
PFDTC	Date/Time of Specimen Collection	Char	ISO 8601	Timing	Date/time of specimen collection	Exp

PFDY	Study Day of Specimen Collection	Num		Timing	1. Study day of specimen collection, measured as integer days. 2. Algorithm for calculations must be relative to the sponsor-defined RFSTDTC variable in Demographics. This formula should be consistent across the submission.	Perm
PFTPT	Planned Time Point Name	Char		Timing	1. Text Description of time when specimen should be taken. 2. This may be represented as an elapsed time relative to a fixed reference point, such as time of last dose. See PFTPTNUM and PFTPTREF. Examples**:** Start, 5 min post.	Perm
PFTPTNUM	Planned Time Point Number	Num		Timing	Numerical version of PFTPT to aid in sorting.	Perm
PFELTM	Elapsed Time from Reference Point	Char	ISO 8601	Timing	Elapsed time (in ISO 8601) relative to a planned fixed reference (PFTPTREF). This variable is useful where there are repetitive measures. Not a clock time or a date time variable. Examples**:** '-P15M' to represent the period of 15 minutes prior to the reference point indicated by PFTPTREF, or 'P8H' to represent the period of 8 hours after the reference point indicated by PFTPTREF.	Perm
PFTPTREF	Time Point Reference	Char		Timing	Name of the fixed reference point referred to by PFELTM, PFTPTNUM, and PFTPT. Examples**:** PREVIOUS DOSE, PREVIOUS MEAL.	Perm
PFRFTDTC	Date/Time of Reference Time Point	Char	ISO 8601	Timing	Date/time of the reference time point, PFTPTREF.	Perm

** Indicates variable may be subject to controlled terminology, (Parenthesis indicates CDISC/NCI codelist code value)*

4.1.2 Assumptions for Pharmacogenomics Findings domain model

1. PF captures results for genetic variation and gene expression. Genetic variation results are indicated by a PFCAT value of GENETIC VARIATION. Gene expression results are indicated by a PFCAT value of GENE EXPRESSION.
2. This domain is for clinical and pre-clinical use, for tests on a study subject or on an infectious microbe.
3. PFRUNID is used to distinguish between records for the same genetic test performed using different assays. When a genetic test is performed on an individual subject using multiple assays, the combination of PFNAM, PFRUNID and REFID may be necessary to obtain the full set of genomic data produced and sent by the lab for a specific test. This can facilitate delivering additional information to regulatory agencies, if needed.
4. PFMETHOD lists techniques for the execution of genomics or genetic testing.
5. Somatic/germline is determined by where the sample came from. Mutations found in tumor tissue but not found in a sample part of the body not affected by cancer (often using a check scraping) are considered somatic. Mutations found in the “control” sample from a non-tumor part of the body are considered germline.
6. Calculations:
    a. When the sponsor performs a calculation, the value for PFDRVFL should be set to “Y”. If PFDRVFL is null, then any calculations have been performed by the vendor.
    b. Only the p-value calculations that are performed by the lab should be included in PF.

7. Viral Findings:
    a. When using the pharmacogenomics domains for viral test reporting, the identification of the virus requires the virus name be placed in the PFSPCIES variable and if available, the strain is placed in PFSTRAIN field. PFSPCIES and PFSTRAIN should not be used for the species and strain of study subjects.
    b. Because viruses only have inheritable DNA, for viral findings, mutation type (PFMUTYP) should always be "GERMLINE".
8. Genetic Variation:
    a. Unless a test is looking for a specific mutation, PFTESTCD and PFTEST generally specify the type of material assessed, such as nucleic acid, amino acid, or codon.
    b. PFSCAT is used to categorize the tests, for example, AMINO ACID, MUTATION, and IDENTIFIER.
    c. Results in PFSTRESC use HGVS nomenclature.
        i. Results from sequencing of the whole gene include the gene's reference sequence accession number. Example: NM_1234c.567A>C
        ii. Results of targeted tests report only the targeted change, not including the initial code that indicates the gene. Example: c.567A>C
    d. When results indicate a mixture of genetic results, as when comparing chromosomes, the results in PFORRES should be concatenated using slashes. For example, "C/T" at a nucleotide position indicates that the subject has heterozygous alleles.

4.1.3 Examples for Pharmacogenomics Findings Domain Model

This copy of the PF domain contains an abbreviated number of examples.

Example 1
This example shows a submission for viral genetics which reports amino acid observations only, without identifying the nucleotide variation that caused it.

Row	STUDYID	DOMAIN	USUBJID	PFSEQ	PFREFID	PFGENTYP	PFGENRI	PFTESTCD	PFTEST	PFSPCIES	PFSTRAIN	PFCAT	PGGENLOC
1	P7081-5101	PF	P7081-5101-01201	1	ABC-001	PROTEIN	NS5B	AA	Amino Acid	HCV	1a	GENETIC VARIATION	65

Row	PFORRES	PFORREF	PFSTRESC	PFREFSEQ	PFRESCAT	PFMUTYP	PFNAM	PFMETHOD	PFBLFL	VISITNUM	VISIT	VISITDY	PFDTC
1 (cont)	R	Gln	NC_00472.1:p.Gln65Arg	NC_00472.1	Point Mutation	GERMLINE	Acme Genetics	Massively Parallel Sequencing	Y	1	Baseline	1	2003-03-27

Example 2
This viral genetics example contains both amino acid observations and their underlying mutations being reported at the codon level.

Row 2: Shows the amino acid that resulted from the mutation.
Row 3: Reports the genetic mutation at the codon level. Because HGVS nomenclature does not support reporting at the codon level, PFSTRESC holds the nucleotide variation.

Row	STUDYID	DOMAIN	USUBJID	PFSEQ	PFGRPID	PFREFID	GENTYP	PFGENRI	PFTESTCD	PFTEST	PFSPCIES	PFSTRAIN	PFCAT	PFGENLOC
1	P7081-5101	PF	P7081-5101-06891	5	NS5B_71	ABC-003	PROTEIN	NS5B	AA	Amino Acid	HCV	1a	GENETIC VARIATION	71
2	P7081-5101	PF	P7081-5101-06891	7	NS5B_71	ABC-003	PROTEIN	NS5B	CDN	Codon	HCV	1a	GENETIC VARIATION	213_215

Row	PFORRES	PFSTRESC	PFORREF	PFREFSEQ	PFNAM	PFSPEC	PFMETHOD	PFBLFL	PFDRVFL	VISITNUM	VISIT	VISITDY	PFDTC
1 (cont)	Ile	NC_00472.1:p.Val71Ile	Val	NC_00472.1	Acme Genetics	DNA	Pyrosequencing	Y	Y	1	Baseline	1	2003-03-27
2 (cont)	ATT	NM_005678:c.213A>G	GTT	NC_00472.1	Acme Genetics	DNA	Pyrosequencing	Y		1	Baseline	1	2003-03-27

Example 3

A frame shift occurs when an insertion or deletion (or both) occurs between the first (*initiation* (ATG)) and last (*termination, stop* (TAA, TAG, or TGA)) codon, replacing the normal C-terminal sequence with one encoded by another reading frame[*]. The effect can be akin to a phrase that reads: "THE OGA TET HER ATF ORT HEC ATW AST OOF AT" because the letter D has been eliminated from the fourth position[†]. Frame shifts can also drastically alter the length of the coding sequence.

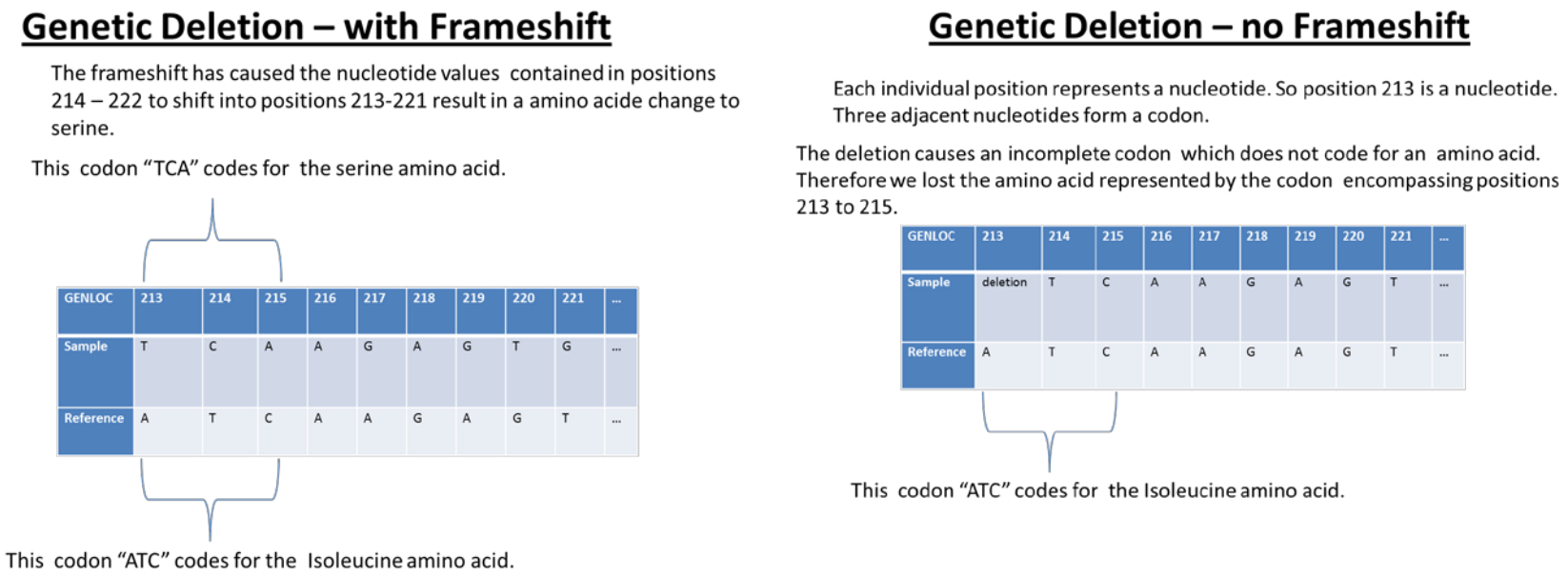


Figure 1: Frame Shift Comparison

[*] den Dunnen J. Recommendations for the description of protein sequence variants (v2.0). *Human Genome Variation Society*. October 11, 2013. Available at: http://www.hgvs.org/mutnomen/recs-prot.html. Accessed December 9, 2013.

[†] The proper phrase, from which D has not been deleted in position 4, would be: "THE DOG ATE THE RAT FOR THE CAT WAS TOO FAT".

This example shows a frame shift found in a virus, and one found in a study subject. Note that in the first two rows, which hold viral genetic data, PFSPCIES and PFSTRAIN have been populated with the virus's species and strain. The second two rows hold data for the the subject, and so PFSPCIES and PFSTRAIN are null.

Row 2: Shows a frame shift itself discovered in the HCV 1a virus carried by subject P341-5101-06345.
Row 1: Shows the nucleotide deletion which caused the frame shift.
Rows 3-4: Show a frame shift and its cause for subject P341-5101-06345, the virus's host.

pf.xpt

Row	STUDYID	DOMAIN	USUBJID	PFSEQ	PFREFID	PFGENTYP	PFGENRI	PFTESTCD	PFTEST	PFSPCIES	PFSTRAIN	…	PFGENLOC	PFORRES	PFORREF
1	P7081	PF	P7081-06345	2	ABC-004	PROTEIN	NS5B	AA	Amino Acid	HCV	1a	…	71	Ser	Ile
2	P7081	PF	P7081-06345	1	ABC-004	PROTEIN	NS5B	NUC	Nucleotide	HCV	1a	…	213	-	A
3	P7081	PF	P7081-06345	1	ABC-009	GENE	CYP3A5	AA	Amino Acid			…	152	Gly	Trp
4	P7081	PF	P7081-06345	2	ABC-009	GENE	CYP3A5	NUC	Nucleotide			…	454	-	T

Row	PFSTRESC	PFREFSEQ	PFREFSEQ	PFRCVNM	PFRESCAT	PFNAM	PFMETHOD	PFRUNID	PFBLFL	VISITNUM	PFDTC
1 (cont)	NP_001234:p.Ile71Serfs	NP_001234	rs1110222	dbSNP	Frame Shift	Acme Genetics	Massively Parallel Sequencing	D391395-001	Y	1	2014-04-01
2 (cont)	NM_005678:c.213delA	NM_005678	rs1110222	dbSNP	Deletion	Acme Genetics	Massively Parallel Sequencing	D391395-001	Y	1	2014-04-01
3 (cont)	NC_000097.3:p.Trp152Glyfs	NP_000097.3	rs5030655	dbSNP	Frame Shift	Acme Genetics	Massively Parallel Sequencing	D391395-005	Y	1	2014-04-01
4 (cont)	NM_000106.5:c.454delT	NM_000106.5	rs5030655	dbSNP	Deletion	Acme Genetics	Massively Parallel Sequencing	D391395-005	Y	1	2014-04-01

Example 4

This example shows nucleotide reads for two subjects using Wave Sequencing. While HGVS recommends that all variations for an individual be given in a single line, the PF domain holds one record per result. Therefore, the results in PFSTRESC, though in HGVS nomenclature, are recorded independent of each other, as if each variation were the only variation observed. When no difference from the reference sequence has been found in the exon specified by PFGENSRI, NVP (for "no variant present") is used in the PFORRES variable; HGVS represents this with an equals sign (=).

Rows 1-7: Show the variations found in first subject's TP53 gene. Note that there are two records for exon 4 because two variations were found.
Row 8: Shows the variation found in exon 2 in the first subject's KRAS gene.
Rows 8-14: Show the variations found in the second subject's TP53 gene. Note that no variations were found in exon 5 (row 10).
Row 15: Shows that no variations were found in exon 2 of the second subject's KRAS gene.

Row	STUDYID	DOMAIN	USUBJID	PFSEQ	PFREFID	PFTESTCD	PFTEST	PFGENTYP	PFGENRI	PFGENSRI	PFCAT	PFGENLOC
1	ABCC	PF	ABCC-XY1	1	AA481093-21-XXX1	NUC	NUCLEOTIDE	GENE	TP53	EXON 4	GENETIC VARIATION	215
2	ABCC	PF	ABCC-XY1	2	AA481093-21-XXX1	NUC	NUCLEOTIDE	GENE	TP53	EXON 4	GENETIC VARIATION	155_156
3	ABCC	PF	ABCC-XY1	3	AA481093-21-XXX1	NUC	NUCLEOTIDE	GENE	TP53	EXON 5	GENETIC VARIATION	469
4	ABCC	PF	ABCC-XY1	4	AA481093-21-XXX1	NUC	NUCLEOTIDE	GENE	TP53	EXON 6	GENETIC VARIATION	639
5	ABCC	PF	ABCC-XY1	5	AA481093-21-XXX1	NUC	NUCLEOTIDE	GENE	TP53	EXON 7	GENETIC VARIATION	700_702
6	ABCC	PF	ABCC-XY1	6	AA481093-21-XXX1	NUC	NUCLEOTIDE	GENE	TP53	EXON 8	GENETIC VARIATION	673-1
7	ABCC	PF	ABCC-XY1	7	AA481093-21-XXX1	NUC	NUCLEOTIDE	GENE	TP53	EXON 9	GENETIC VARIATION	993+12
8	ABCC	PF	ABCC-XY1	8	AA481093-21-XXX1	NUC	NUCLEOTIDE	GENE	KRAS	EXON 2	GENETIC VARIATION	34
9	ABCC	PF	ABCC-XY2	1	AA811854-21-XXX2	NUC	NUCLEOTIDE	GENE	TP53	EXON 4	GENETIC VARIATION	188_194
10	ABCC	PF	ABCC-XY2	2	AA811854-21-XXX2	NUC	NUCLEOTIDE	GENE	TP53	EXON 5	GENETIC VARIATION	
11	ABCC	PF	ABCC-XY2	3	AA811854-21-XXX2	NUC	NUCLEOTIDE	GENE	TP53	EXON 6	GENETIC VARIATION	639

Row	STUDYID	DOMAIN	USUBJID	PFSEQ	PFREFID	PFTESTCD	PFTEST	PFGENTYP	PFGENRI	PFGENSRI	PFCAT	PFGENLOC
12	ABCC	PF	ABCC-XY2	4	AA811854-21-XXX2	NUC	NUCLEOTIDE	GENE	TP53	EXON 7	GENETIC VARIATION	743
13	ABCC	PF	ABCC-XY2	5	AA811854-21-XXX2	NUC	NUCLEOTIDE	GENE	TP53	EXON 8	GENETIC VARIATION	873
14	ABCC	PF	ABCC-XY2	6	AA811854-21-XXX2	NUC	NUCLEOTIDE	GENE	TP53	EXON 9	GENETIC VARIATION	926
15	ABCC	PF	ABCC-XY2	7	AA811854-21-XXX2	NUC	NUCLEOTIDE	GENE	KRAS	EXON 2	GENETIC VARIATION	

Row	PFORRES	PFSTRESC	PFORREF	PFREFSEQ	PFRESCV	PFRCVNM	PFNAM	PFSPEC	PFMETHOD	VISITNUM	VISIT	PFDTC
1 (cont)	G	NM_000546.5:c.215C>G	C	NM_000546.5	rs1042522	dbSNP	Vendor A	DNA	WAVE-Sequencing	-1	Pre-Treatment	2012-04-01
2 (cont)	-	NM_000546.5:c.155_156del	AA	NM_000546.5			Vendor A	DNA	WAVE-Sequencing	-1	Pre-Treatment	2012-04-01
3 (cont)	T	NM_000546.5:c.469G>T	G	NM_000546.5			Vendor A	DNA	WAVE-Sequencing	-1	Pre-Treatment	2012-04-01
4 (cont)	G	NM_000546.5:c.639A>G	A	NM_000546.5	rs1800572	dbSNP	Vendor A	DNA	WAVE-Sequencing	-1	Pre-Treatment	2012-04-01
5 (cont)	-	NM_000546.5:c.700_702del	TAC	NM_000546.5			Vendor A	DNA	WAVE-Sequencing	-1	Pre-Treatment	2012-04-01
6 (cont)	C	NM_000546.5:c.673-1G>C	G	NM_000546.5			Vendor A	DNA	WAVE-Sequencing	-1	Pre-Treatment	2012-04-01
7 (cont)	C	NM_000546.5:c.993+12T>C	T	NM_000546.5	rs1800899	dbSNP	Vendor A	DNA	WAVE-Sequencing	-1	Pre-Treatment	2012-04-01
8 (cont)	T	NM_033360.2 :c.34G>T	G	NM_033360.2			Vendor A	DNA	WAVE-Sequencing	-1	Pre-Treatment	2012-04-01
9 (cont)	-	NM_000546.5:c.188_194del	CTCCCAG	NM_000546.5			Vendor A	DNA	WAVE-Sequencing	-1	Pre-Treatment	2012-04-01
10 (cont)	NVP	NM_000546.5:c.=		NM_000546.5			Vendor A	DNA	WAVE-Sequencing	-1	Pre-Treatment	2012-04-01
11 (cont)	G	NM_000546.5:c.639A>G	A	NM_000546.5	rs1800372	dbSNP	Vendor A	DNA	WAVE-Sequencing	-1	Pre-Treatment	2012-04-01
12 (cont)	A	NM_000546.5:c.743G>A	G	NM_000546.5	rs11540652	dbSNP	Vendor A	DNA	WAVE-Sequencing	-1	Pre-Treatment	2012-04-01
13 (cont)	-	NM_000546.5:c.873delG	G	NM_000546.5			Vendor A	DNA	WAVE-Sequencing	-1	Pre-Treatment	2012-04-01
14 (cont)	-	NM_000546.5:c.926delC	C	NM_000546.5			Vendor A	DNA	WAVE-Sequencing	-1	Pre-Treatment	2012-04-01
15 (cont)	NVP	NM_033360.2 :c.=		NM_033360.2			Vendor A	DNA	WAVE-Sequencing	-1	Pre-Treatment	2012-04-01

Example 5

A polymerase chain reaction (PCR) is a common method for amplifying DNA. Quantitative PCR and DNA microarray are modern methodologies for studying gene expression. The amount of expressed gene in a cell can be measured by the number of copies of mRNA transcript of that gene present in a sample. In mRNA-based PCR, the RNA sample is first reverse-transcribed to cDNA. Then the cDNA is quantified via conventional PCR in a thermal cycler. This method is referred to as qRT-PCR (quantitative reverse-transcriptase polymerase chain reaction). The rate of generation of the amplified product is measured at each PCR cycle and the number of cycles required to reach a defined signal threshold is referred to as cycle time (Ct). The data generated can be analyzed by computer software to calculate relative gene expression (mRNA copy number) in several samples which is actually the measurement of the quantity of cDNA produced by the reverse-transcriptase step of the assay.

Rows 1, 7: Report on the copy number.
Rows 4-6, 10-12: Show the raw Ct values.
Rows 3, 9: Show the mean Ct value.

Row	STUDYID	DOMAIN	USUBJID	PFSEQ	PFREFID	PFTESTCD	PFTEST	REPNUM	PFGENTYP	PFGENRI	PFGENSRI	PFCAT
1	AAA	PF	AAA-XXX1	1	L1046058-3-S0072248	COPYNUM	Copy Number		GENE	MYC	EXON Boundary 2-3	GENE EXPRESSION
4	AAA	PF	AAA-XXX1	4	L1046058-3-S0072248	RAWCT	Raw Ct Value	1	GENE	MYC	EXON Boundary 2-3	GENE EXPRESSION
5	AAA	PF	AAA-XXX1	5	L1046058-3-S0072248	RAWCT	Raw Ct Value	2	GENE	MYC	EXON Boundary 2-3	GENE EXPRESSION
6	AAA	PF	AAA-XXX1	6	L1046058-3-S0072248	RAWCT	Raw Ct Value	3	GENE	MYC	EXON Boundary 2-3	GENE EXPRESSION
3	AAA	PF	AAA-XXX1	3	L1046058-3-S0072248	MEANCT	Mean Ct Value		GENE	MYC	EXON Boundary 2-3	GENE EXPRESSION
7	AAA	PF	AAA-XXX1	7	L1046058-3-S0072248	COPYNUM	Copy Number		GENE	MFNG	EXON Boundary 4-5	GENE EXPRESSION
10	AAA	PF	AAA-XXX1	10	L1046058-3-S0072248	RAWCT	Raw Ct Value	1	GENE	MFNG	EXON Boundary 4-5	GENE EXPRESSION
11	AAA	PF	AAA-XXX1	11	L1046058-3-S0072248	RAWCT	Raw Ct Value	2	GENE	MFNG	EXON Boundary 4-5	GENE EXPRESSION

Row	STUDYID	DOMAIN	USUBJID	PFSEQ	PFREFID	PFTESTCD	PFTEST	REPNUM	PFGENTYP	PFGENRI	PFGENSRI	PFCAT
12	AAA	PF	AAA-XXX1	12	L1046058-3-S0072248	RAWCT	Raw Ct Value	3	GENE	MFNG	EXON Boundary 4-5	GENE EXPRESSION
9	AAA	PF	AAA-XXX1	9	L1046058-3-S0072248	MEANCT	Mean Ct Value		GENE	MFNG	EXON Boundary 4-5	GENE EXPRESSION

Row	PFORRES	PFORRESU	PFSTRESC	PFSTRESN	PFSTRESU	PFNAM	PFSPEC	PFMETHOD	PFRUNID	VISIT	VISITNUM	PFDTC
1	34.07	Copies/ng	34.07	34.07	Copies/ng	Vendor B	RNA	qRT-PCR	Hs00153408_m1	Pre-Treatment	-1	2013:01:30:19:56:00
4	26.595144	Cycles	26.595144	26.595144	Cycles	Vendor B	RNA	qRT-PCR	Hs00153408_m1	Pre-Treatment	-1	2013:01:30:19:56:00
5	26.098644	Cycles	26.098644	26.098644	Cycles	Vendor B	RNA	qRT-PCR	Hs00153408_m1	Pre-Treatment	-1	2013:01:30:19:56:00
6	26.399672	Cycles	26.399672	26.399672	Cycles	Vendor B	RNA	qRT-PCR	Hs00153408_m1	Pre-Treatment	-1	2013:01:30:19:56:00
3	26.364487	Cycles	26.364487	26.364487	Cycles	Vendor B	RNA	qRT-PCR	Hs00153408_m1	Pre-Treatment	-1	2013:01:30:19:56:00
7	0.67	Copies/ng	0.67	0.67	Copies/ng	Vendor B	RNA	qRT-PCR	Hs00159117_m1	Pre-Treatment	-1	2013:02:30:19:56:00
10	32.979595	Cycles	32.979595	32.979595	Cycles	Vendor B	RNA	qRT-PCR	Hs00159117_m1	Pre-Treatment	-1	2013:02:30:19:56:00
11	33.54182	Cycles	33.54182	33.54182	Cycles	Vendor B	RNA	qRT-PCR	Hs00159117_m1	Pre-Treatment	-1	2013:02:30:19:56:00
12	33.4799	Cycles	33.4799	33.4799	Cycles	Vendor B	RNA	qRT-PCR	Hs00159117_m1	Pre-Treatment	-1	2013:02:30:19:56:00
9	33.33377	Cycles	33.33377	33.33377	Cycles	Vendor B	RNA	qRT-PCR	Hs00159117_m1	Pre-Treatment	-1	2013:02:30:19:56:00

Example 6

This example contains gene expression data obtained using a two-color (two-channel) microarray. A microarray is a type of lab-on-a-chip, in which the results initially appear as a field of dots of varying intensity and hue, as in the figure below. Single-color (single-channel) microarrays are also available; the results they produce are similar, but vary only in intensity.

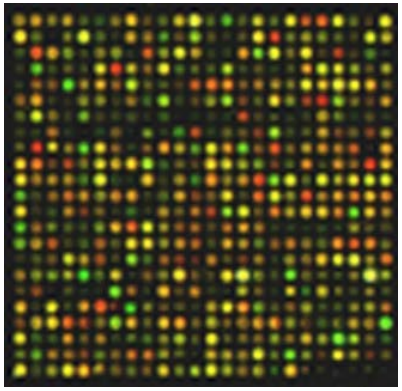


Figure 2: Sample Two-Color Microarray[‡]

Interpretation of the microarray requires a specialized scanner which converts the color and brightness of each dot into numeric intensity values. Some scanners normalize the values as part of the process. Because a single microarray can hold thousands of results, a vendor may choose to report only the "interesting" results (e.g. wells that showed over-expression), which can be determined with specialized analytic software, often sold by the same company that produced the microarray chip and/or the scanner. In this example, AGM-G4851B is the SPDEVID of the microarray kit, AGS- G4900DA is the SPDEVID of the scanner, and

[‡] Public domain. Originally released by the National Cancer Institute.

MANAN03 is the SPDEVID of the workstation to which the scanner is connected, and on which the vendor's analytic software is installed. Because the analysis was performed by the vendor and not by the sponsor, there is no derived value flag.

Row 1: Shows the target sequence being tested. Because the sequence is 60 characters long, for the purposes of this example it has been omitted from PFSTRESC.
Rows 2-3: Show the intensity values for each channel, or color.
Row 4: Shows the p-value.
Row 5: Shows the fold change.

Row	STUDYID	DOMAIN	USUBJID	SPDEVID	PFSEQ	PFGRPID	PFREFID	PFTESTCD	PFTEST	PFCAT	PFORRES
1	A12345	PF	43871	AGM-G4851B	1	1	2287.09443	ACTSEQ	Active Sequence	Pre-Analytic	GGGAGAGAAGAGACCTGCCA GATTATCAGACCTCTTCATG TTAAAAGACCATCTCCTGTA
2	A12345	PF	43871	AGS- G4900DA	2	1	2287.09443	NINT1VAL	Normalized Intensity 1 Value	Pre-Analytic	1.16279
3	A12345	PF	43871	AGS- G4900DA	3	1	2287.09443	NINT2VAL	Normalized Intensity 2 Value	Pre-Analytic	0.96469
4	A12345	PF	43871	MANAN03	4	1	2287.09443	PVAL	P Value	Analytic	0.05391
5	A12345	PF	43871	MANAN03	5	1	2287.09443	FOLDCHG	Fold Change	Analytic	1.8

Row	PFSTRESC	PFSTRESN	PFXFN	PFNAM	PFSPEC	PFMETHOD	PFRUNID	PFANMETH	PFBLFL	VISITNUM	PFDTC
1 (cont)	…		2.16.090.1.135764.3.4:7280912	Deluxe Central Labs	RNA	Microarray	1000450001			2	2005-03-21T11:28:17
2 (cont)	1.16279	1.16279	2.16.090.1.135764.3.4:7280912	Deluxe Central Labs	RNA	Microarray	1000450001	LOWESS		2	2005-03-21T11:28:17
3 (cont)	0.96469	0.96469	2.16.090.1.135764.3.4:7280912	Deluxe Central Labs	RNA	Microarray	1000450001	LOWESS		2	2005-03-21T11:28:17
4 (cont)	0.05391	0.05391	2.16.090.1.135764.3.4:7280912	Deluxe Central Labs	RNA	Microarray	1000450001			2	2005-03-21T11:28:17
5 (cont)	1.8	1.8	2.16.090.1.135764.3.4:7280912	Deluxe Central Labs	RNA	Microarray	1000450001			2	2005-03-21T11:28:17